

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2000, 08:16:18 ; Search time 38.4 Seconds
(without alignments)
244.263 Million cell updates/sec

Title: US-08-900-220-17

Perfect score: 2088
Sequence: 1 MALLNLPLPLCLALLALPA.....PTGMWYSRLYLABELLG 396

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2088	100.0	396	1 W97764	Human Desert hedge
2	2088	100.0	396	1 Y05517	Human Desert hedge
3	2083	99.8	396	1 W97955	Human Desert hedge
4	2008	96.2	396	1 R77345	Mouse desert hedge
5	2008	96.2	396	1 W94469	Mouse Dh hedgehog
6	2008	96.2	396	1 W97766	Mouse Desert hedge
7	2008	96.2	396	1 Y05511	Mouse Desert hedge
8	1974	84.5	396	1 W97958	Human Desert hedge
9	1217	58.3	437	1 R77339	Mouse sonic hedgehog
10	1217	58.3	437	1 W94471	Mouse Sh hedgehog
11	1217	58.3	437	1 W97768	Mouse sonic hedgehog
12	1217	58.3	437	1 Y05513	Mouse sonic hedgehog
13	1214	58.1	437	1 W61488	Mouse sonic hedgehog
14	1198	57.4	425	1 R77338	Chicken sonic hedgehog
15	1198	57.4	425	1 W94468	Chicken Sh hedgehog
16	1198	57.4	425	1 W97765	Chicken Sonic hedgehog
17	1198	57.4	425	1 Y05510	Chicken Sonic hedgehog
18	1177	56.4	425	1 R77341	Human sonic hedgehog
19	1177	56.4	425	1 W94473	Human Sh hedgehog
20	1177	56.4	425	1 W97770	Human sonic hedgehog
21	1177	56.4	425	1 Y05515	Human Sonic hedgehog
22	1173	56.2	425	1 W48736	Human Sonic hedgehog
23	1171.5	56.1	411	1 W94474	Human Ih hedgehog
24	1171.5	56.1	411	1 W97763	Human Ih hedgehog
25	1171.5	56.1	411	1 Y05516	Human Indian hedgehog
26	1171	56.1	462	1 W48735	Human mutated sonic hedgehog
27	1167	55.9	437	1 R80071	Rat Vhh-1, Nucleic acid
28	1158	55.5	425	1 W61487	Chicken sonic hedgehog
29	1154	55.3	416	1 R97658	Zebrafish liggy-wi
30	1154	55.3	416	1 W61485	Zebrafish liggy-wi
31	1154	55.3	416	1 W94475	Zebrafish rh hedgehog
32	1154	55.3	416	1 W97771	Zebrafish liggy-wi
33	1154	55.3	416	1 Y05518	Zebrafish liggy-wi
34	1126	53.9	411	1 W94470	Mouse Ih hedgehog

35	1126	53.9	411	1 W97767	Mouse Indian hedgehog
36	1126	53.9	411	1 Y05512	Mouse Indian hedgehog
37	1120	53.6	418	1 W61486	Zebrafish sonic hedgehog
38	1069	51.2	418	1 R77340	Zebrafish sonic hedgehog
39	1069	51.2	418	1 W94472	Zebrafish Sh hedgehog
40	1069	51.2	418	1 W97769	Zebrafish Sonic hedgehog
41	1069	51.2	418	1 Y05514	Zebrafish Sonic hedgehog
42	1045	50.0	198	1 W79597	Human Desert hedgehog
43	989	47.4	191	1 W79598	Human Desert hedgehog
44	964	46.2	182	1 W79596	Human Desert hedgehog
45	956.5	45.8	336	1 R77343	Mouse Indian hedgehog

ALIGNMENTS

RESULT 1
ID W97764
AC W97764; standard; Protein; 396 AA.
DT 21-MAY-1999 (first entry)
DE Human Desert hedgehog (Dhh) protein.
KW Sesert hedgehog; Dhh protein; human; dopaminergic; GABA-ergic; ptc therapeutic; patched; signal transduction; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW cerebral ischaemia; hypoxia; neuroprotective; therapy.
OS Homo sapiens.
PN W09904775-A2.
PD 04-FEB-1999.
PE 24-JUL-1998; U15419.
PR 24-JUL-1997; US-900220.
PI (ONTO-) ONTOGENY INC.
PI Mahanahappa NK, Miao N, Pang K, Wang M;
DR WPI: 99-142578/12.
NR N-PSDB: X07270.
PT Increasing the survival of neuronal, dopaminergic and GABA-ergic cells - by using a ptc therapeutic such as a protein kinase inhibitor, or an agent derived from hedgehog polypeptides, useful in the treatment of Parkinson's disease.
PS Claim 30: Page 97-99; 138pp; English.
CC This polypeptide is human Dhh Desert hedgehog protein. The invention is based on the finding that hedgehog proteins are useful as protective agents in the treatment and prophylaxis of neurodegenerative disorders resulting from the loss of dopaminergic and/or GABA-ergic neurons, or the general loss of tissue from the substantia nigra. Exemplary disorders include Parkinson's disease, Huntington's disease (both claimed), amyotrophic lateral sclerosis and cerebral ischaemia. The invention relates to hedgehog therapeutics (i.e. hedgehog polypeptides and gene therapy constructs e.g. constructs encoding recombinant hedgehog polypeptides and trans-activation constructs for altering hedgehog gene regulatory sequences) and ptc therapeutics (i.e. agents which mimic the effect of naturally occurring hedgehog proteins on patched signalling) that are effective in both human and animal subjects. A bioactive polypeptide comprising amino acids 23-198 of human Dhh is preferred. The products can also be used for the maintenance of differentiated neurons in cultures, and to enhance the implantation of such neuronal cells in an animal. They can be used to prevent or treat neurodegenerative conditions arising from the use of certain drugs, and in the prevention and/or treatment of hypoxia, e.g. as a neuroprotective agent.
SQ Sequence 396 AA;

Query Match 100.0%; Score 2088; DB 1; Length 396;
Best Local Similarity 100.0%; Pred. No. 1.1e-216;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MALLNLPLPLCLALLALPAOSCGRPVGRARRAROLVLYLKQVPVPTLGAS 60
DB 1 MALLNLPLPLCLALLALPAOSCGRPVGRARRAROLVLYLKQVPVPTLGAS 60
OY 61 GPAEGRVARGSERFDLVPNPNPDIIFKDENSGADRLMTERCKERNALAIAYVMMPG 120

```

DB 61 GAEGEVARSERFRDLVFNYPNDIIIFKDENSGADRLMTERCKERNALAIAYNNMMPG 120
QY 121 VRLRVTEGMEDEGHHADSLHYEGRALDITTSDRDNKRYGLRLARLAVEAGFDMVYYESRN 180
DB 121 VRLRVTEGMEDEGHHADSLHYEGRALDITTSDRDNKRYGLRLARLAVEAGFDMVYYESRN 180
QY 181 HHVSVKADNSLAVRAGGCEPFGNATVRLMSGERRKGLRELHRDWMVLAADASGRVPTPV 240
DB 181 HHVSVKADNSLAVRAGGCEPFGNATVRLMSGERRKGLRELHRDWMVLAADASGRVPTPV 240
QY 241 LFLDRDLQRASFAVAVETEMPPRKLLTPMHLVFAARGPAPGDFAPVFAARLRAGDSV 300
DB 241 LFLDRDLQRASFAVAVETEMPPRKLLTPMHLVFAARGPAPGDFAPVFAARLRAGDSV 300
QY 301 LAPGDALRPARVAVAREEAVGVFAPLTAHGTLLVNDVLASCYAVLESQMAHRAFAPL 360
DB 301 LAPGDALRPARVAVAREEAVGVFAPLTAHGTLLVNDVLASCYAVLESQMAHRAFAPL 360
QY 361 RLHLALGALLPGGAVOPTGMHWSRLLYRLAEELG 396
DB 361 RLHLALGALLPGGAVOPTGMHWSRLLYRLAEELG 396

RESULT 2
ID Y05517 standard; Protein: 396 AA.
AC Y05517;
DE 05-JUL-1999 (first entry)
DE Human Desert hedgehog protein Dhh.
KW Desert hedgehog; Dhh protein; human; hedgehog therapeutic;
KW ptc therapeutic; patched; signal transduction; muscle atrophy;
KW cachaexia; muscular myopathy; myoblastic sarcoma; therapy.
OS Homo sapiens.
PN W09910004-A2.
PD 24-MAR-1999.
PF 28-AUG-1998; U17922.
PR 29-AUG-1997; US-057394.
PA (ONTO-) ONTOGENY INC.
PI 3ladgen CS, Currie PD, Hughes SM, Ingham PW;
DR WPI: 99-243557/20.
DR N-PSDB: X25105.
PT A new method to regulate muscle growth
PS Disclosure; page 121-121; 130pp; English.
CC The present sequence is human Desert hedgehog protein Dhh. The
CC invention relates to a method for modulating the formation and/or
CC maintenance of muscle tissue by ecotopically contacting muscle
CC cells, especially muscle stem/progenitor cells, in vitro or in
CC vivo, with a hedgehog therapeutic (i.e. hedgehog polypeptides and
CC gene therapy constructs) or ptc therapeutic (i.e. a small organic
CC molecule that mimics the effect of hedgehog proteins on patched
CC signaling, or activates or potentiates patched signaling) in an
CC amount effective to alter the growth state of the treated cells.
CC Also claimed is a method for treatment or prevention of disorders
CC of, or surgical or cosmetic repair of, such muscle tissues, by
CC administering a hedgehog polypeptide or ptc therapeutic. The
CC disorder may be muscle atrophy, in particular skeletal muscle
CC atrophy or cardiac muscle atrophy, cachaexia, or muscular myopathy
CC (all claimed). The hedgehog polypeptide or ptc therapeutic can
CC inhibit growth of myoblastic-derived tissue to provide treatment of
CC hyperblastic or neoplastic growth of muscle tissue such as in
CC myoblastic sarcoma (also claimed). The hedgehog therapeutic
CC preferably comprises at least a bioactive extracellular portion of
CC a hedgehog protein (see Y05510-19) encoded by a vertebrate hedgehog
CC gene (see X25098-107), especially a human hedgehog gene.
SO Sequence 396 AA.

```

```

Query Match 100.0%; Score 2088; DB 1; Length 396;
Best Local Similarity 100.0%; Pred. No. 1.1e-216;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALLTNLPCLCLALLALPAQSCGPGRGVGRRRYARKOLVPLLYKQFVPGVPTLGAS 60

```

```

DB 61 GAEGEVARSERFRDLVFNYPNDIIIFKDENSGADRLMTERCKERNALAIAYNNMMPG 120
QY 61 GAEGEVARSERFRDLVFNYPNDIIIFKDENSGADRLMTERCKERNALAIAYNNMMPG 120
DB 61 GAEGEVARSERFRDLVFNYPNDIIIFKDENSGADRLMTERCKERNALAIAYNNMMPG 120
QY 121 VRLRVTEGMEDEGHHADSLHYEGRALDITTSDRDNKRYGLRLARLAVEAGFDMVYYESRN 180
DB 121 VRLRVTEGMEDEGHHADSLHYEGRALDITTSDRDNKRYGLRLARLAVEAGFDMVYYESRN 180
QY 181 HHVSVKADNSLAVRAGGCEPFGNATVRLMSGERRKGLRELHRDWMVLAADASGRVPTPV 240
DB 181 HHVSVKADNSLAVRAGGCEPFGNATVRLMSGERRKGLRELHRDWMVLAADASGRVPTPV 240
QY 241 LFLDRDLQRASFAVAVETEMPPRKLLTPMHLVFAARGPAPGDFAPVFAARLRAGDSV 300
DB 241 LFLDRDLQRASFAVAVETEMPPRKLLTPMHLVFAARGPAPGDFAPVFAARLRAGDSV 300
QY 301 LAPGDALRPARVAVAREEAVGVFAPLTAHGTLLVNDVLASCYAVLESQMAHRAFAPL 360
DB 301 LAPGDALRPARVAVAREEAVGVFAPLTAHGTLLVNDVLASCYAVLESQMAHRAFAPL 360
QY 361 RLHLALGALLPGGAVOPTGMHWSRLLYRLAEELG 396
DB 361 RLHLALGALLPGGAVOPTGMHWSRLLYRLAEELG 396

RESULT 3
ID W79595 standard; Protein: 396 AA.
AC W79595;
DE 02-FEB-1999 (first entry)
DE Human Desert hedgehog protein precursor.
KW Desert hedgehog; Huhhh; human; monoclonal antibody.
OS Homo sapiens.
PN Key
FT Peptide 1..22
FT Location/Qualifiers
FT /label= Sig_peptide
FT 23..198
FT Protein /label= Mat_protein
PN EP-874048-A2.
PD 28-OCT-1998.
PF 24-APR-1998; 303187.
PR 14-APR-1998; JP-117873.
PR 25-APR-1997; JP-121578.
PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
PI Ariyasu T, Nakamura S, Orita K;
DR WPI: 98-544642/47.
DR N-PSDB: V62395.
PT Human Desert hedgehog protein - and corresponding DNA and monoclonal
PT antibody
PS Claim 4: Page 24-25; 39pp; English.
CC This is the amino acid sequence of a precursor of a novel human
CC Desert hedgehog protein, as deduced from a cDNA clone (see V62395)
CC derived from cell line ARH-77 (ATCC CRL-1621). The mature form
CC (see W79593) of the hedgehog protein is also claimed, as are a
CC monoclonal antibody (Mab) that recognises the protein, a process
CC for producing the protein, and a method for detecting the protein.
CC The hedgehog protein is useful in establishment of hybridomas which
CC produce antibodies recognising the protein, and the Mab is useful
CC for detecting and purifying the protein. The hedgehog protein,
CC DNA and Mab can be used to elucidate hereditary morphological
CC abnormalities in humans to establish their treatments and diagnoses.
SO Sequence 396 AA.

```

```

Query Match 99.8%; Score 2083; DB 1; Length 396;
Best Local Similarity 99.5%; Pred. No. 3.7e-216;
Matches 394; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MALLTNLPCLCLALLALPAQSCGPGRGVGRRRYARKOLVPLLYKQFVPGVPTLGAS 60

```

Db 1 MALTLNLPICCLALLPAOSCGRGPGVGRRRYARKQVLPVLYKQFVPGVPERTIGAS 60
QY 61 GPAGRVARSGSERPDLVYNNPDIIFKDEENSGADRLMTERCKERNALAIAMNMMPG 120
Db 61 GPAGRVARSGSERPDLVYNNPDIIFKDEENSGADRLMTERCKERNALAIAMNMMPG 120
QY 121 VRLVTEGMDGDEGHHADSLHYEGRALDITTSDDRNKYGGLARLAVAGFDWYYESRN 180
Db 121 VRLVTEGMDGDEGHHADSLHYEGRALDITTSDDRNKYGGLARLAVAGFDWYYESRN 180
QY 181 HHVSVKADNSLAVRAGCGPGNATVRLMSGERGLRELHGGDWVLAADAGRVPVPTVL 240
Db 181 HHVSVKADNSLAVRAGCGPGNATVRLMSGERGLRELHGGDWVLAADAGRVPVPTVL 240
QY 241 LFLDRDLORRASFAVAVETEMPRLKLLTPMHLVFAARGPAPAGDFAPVFARRLRAGDSV 300
Db 241 LFLDRDLORRASFAVAVETEMPRLKLLTPMHLVFAARGPAPAGDFAPVFARRLRAGDSV 300
QY 301 LAPGDLARPARVARVAREEAVGFAPLTAHGTLLVNDVLAACYAVLESHOMAHRAFAPL 360
Db 301 LAPGDLARPARVARVAREEAVGFAPLTAHGTLLVNDVLAACYAVLESHOMAHRAFAPL 360
QY 361 RLHLALGALPGAVOPTGMHWSRLYLRLAEELG 396
Db 361 RLHLALGALPGAVOPTGMHWSRLYLRLAEELG 396

RESULT 4

R77345
ID R77345 standard; Protein; 396 AA.

AC R77345;
DE Mouse desert hedgehog protein.
KW Mouse; desert hedgehog protein; probe; primer; diagnostic;
KM nervous system disorder; gene therapy; antibody.
OS Mus musculus.
FH Key
FT Peptide 1..22
FT /note= "signal peptide"
FT 23..28
FT /note= "conserved sequence (R77349)"
FT peptide
FN W0518856-A1.
PD 13-JUL-1995.
PE 30-DEC-1994; U14992.
PR 30-DEC-1993; US-176427.
PR 14-DEC-1994; US-356060.
PA (HARD) HARVARD COLLEGE.
PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
PI Ingham PW, McMahon AP, Tablin CJ;
DR N-PSDB: Q91642.
PT Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful
PT to treat degenerative nervous system disorder(s) and in gene
PT therapy.
PS Claim 17; Page 135-37; 210pp; English.
CC The sequence represents a mouse Indian hedgehog protein, homologous
CC to a Drosophila hedgehog protein (R77337), and is encoded by a cDNA
CC isolated by low stringency screening of a mouse genome DNA library.
CC probes and primers derived from hedgehog sequences may be used as
CC diagnostic agents for neuromuscular, autonomic or central nervous
CC system disorders, and the gene may also be used in gene therapy.
CC Antibodies generated from the protein may be used as therapeutic or
CC research reagents.
SQ Sequence 396 AA;

Query Match 96.2%; Score 2008; DB 1; Length 396;
Best Local Similarity 96.5%; Pred. No. 4.5e-208;
Matches 382; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 MALTLNLPICCLALLPAOSCGRGPGVGRRRYARKQVLPVLYKQFVPGVPERTIGAS 60
Db 1 MALTLNLPICCLALLPAOSCGRGPGVGRRRYARKQVLPVLYKQFVPGVPERTIGAS 60

QY 61 GPAGRVARSGSERPDLVYNNPDIIFKDEENSGADRLMTERCKERNALAIAMNMMPG 120
Db 61 GPAGRVARSGSERPDLVYNNPDIIFKDEENSGADRLMTERCKERNALAIAMNMMPG 120
QY 121 VRLVTEGMDGDEGHHADSLHYEGRALDITTSDDRNKYGGLARLAVAGFDWYYESRN 180
Db 121 VRLVTEGMDGDEGHHADSLHYEGRALDITTSDDRNKYGGLARLAVAGFDWYYESRN 180
QY 181 HHVSVKADNSLAVRAGCGPGNATVRLMSGERGLRELHGGDWVLAADAGRVPVPTVL 240
Db 181 HHVSVKADNSLAVRAGCGPGNATVRLMSGERGLRELHGGDWVLAADAGRVPVPTVL 240
QY 241 LFLDRDLORRASFAVAVETEMPRLKLLTPMHLVFAARGPAPAGDFAPVFARRLRAGDSV 300
Db 241 LFLDRDLORRASFAVAVETEMPRLKLLTPMHLVFAARGPAPAGDFAPVFARRLRAGDSV 300
QY 301 LAPGDLARPARVARVAREEAVGFAPLTAHGTLLVNDVLAACYAVLESHOMAHRAFAPL 360
Db 301 LAPGDLARPARVARVAREEAVGFAPLTAHGTLLVNDVLAACYAVLESHOMAHRAFAPL 360
QY 361 RLHLALGALPGAVOPTGMHWSRLYLRLAEELG 396
Db 361 RLHLALGALPGAVOPTGMHWSRLYLRLAEELG 396

RESULT 5

W94469
ID W94469 standard; Protein; 396 AA.

AC W94469;
DE Mouse Dhh hedgehog protein sequence.
KW Patched; hedgehog; ptc therapeutic; neuroprotective; neuronal cell;
KM brain infarction; cerebral infarction; transient ischaemic attack;
KM stroke; cerebral infarct volume; spinal cord; oedema; trauma;
KM haemorrhage; encephalomyelitis; coronary bypass; cerebral hypoxia.
OS Mus sp.
FN W09900117-A2.
PD 07-JAN-1999.
PE 26-JUN-1998; U13387.
PR 27-JUN-1997; US-883656.
PA (ONTO-) ONTOGENY INC.
PI Mahanthappa NK;
DR N-PSDB: X16183.
PT Method for limiting damage to neurons caused by ischaemic or epoxic
PT conditions - is used for the treatment and prevention of e.g.
PT cerebral infarction, stroke and transient ischaemic attacks
PS Disclosure; page 65-66; 104pp; English.
CC A method has been developed for limiting the damage to neuronal cells by
CC ischaemic or epoxic conditions by administering a ptc (patched).
CC therapeutic agent to reduce cerebral infarct volume (CIV). Damage to
CC neuronal cells can also be limited by administering a gene activation
CC construct which recombinates with the genomic hedgehog gene to provide a
CC heterologous transcription regulator linked to the coding region of this
CC gene. Administration of the ptc therapeutic agent is used to protect
CC cerebral tissues against ischaemic injury; to treat cerebral infarct or
CC ischaemia, stroke (thrombotic or embolic) and transient ischaemic
CC attacks. It may also be used as a prophylactic in many other cases of
CC injury to the brain or spinal cord, oedema caused by trauma, haemorrhage
CC and encephalomyelitis, or in conjunction with (coronary bypass) surgery.
CC Treatment (which may be prophylactic) is used where ischaemic/epoxic
CC conditions may cause cerebral hypoxia, or progressive loss of neurons
CC due to oxygen depletion, including in patients with hypotension. The
CC treatment reduces CIV by at least 25, particularly at least 70.%. The
CC present sequence represents a hedgehog sequence given in the present
CC invention.
SQ Sequence 396 AA;

Query Match 96.2%; Score 2008; DB 1; Length 396;
Best Local Similarity 96.5%; Pred. No. 4.5e-208;
Matches 382; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 MALLTNLPLCLALLALPAQSCGPGRGVRRRRARAKQVLPPLYKQVPGVPERTLGAS 60
 DB 1 MALPASLPLCLALLALSAQSCGPGRGVRRRRARAKQVLPPLYKQVPGVPERTLGAS 60
 QY 61 GAEGGVARGSERFRDLVNPNDIIFKDEENSGADRLMTERCKERVNALAIVANNMMPG 120
 DB 61 GAEGGVARGSERFRDLVNPNDIIFKDEENSGADRLMTERCKERVNALAIVANNMMPG 120
 QY 121 VALRTGEMDEGHHADSLHYEGRALDITTSDRRNKYGGLARLAVAGDFWYYESRN 180
 DB 121 VALRTGEMDEGHHADSLHYEGRALDITTSDRRNKYGGLARLAVAGDFWYYESRN 180
 QY 181 HHVSVKADNSLAVRAGCGCPGNATVRLMSGERKGLRELHGDWVLAADAGRVVPTVYL 240
 DB 181 HHVSVKADNSLAVRAGCGCPGNATVRLMSGERKGLRELHGDWVLAADAGRVVPTVYL 240
 QY 241 LFLDLDLORRASFVAVETERPRLKLLTPWHLVFARGPAPAGDFAVFARRLRAGDSV 300
 DB 241 LFLDLDLORRASFVAVETERPRLKLLTPWHLVFARGPAPAGDFAVFARRLRAGDSV 300
 QY 301 LAPGSDALPAPVARVAREEAVGFAPLTAHGTLLVNDVLASCVAVLSEHOMAHRAFAPL 360
 DB 301 LAPGSDALPAPVARVAREEAVGFAPLTAHGTLLVNDVLASCVAVLSEHOMAHRAFAPL 360
 QY 361 RLHALGALLPGAGVOPGMHYSRLYRLAEELG 396
 DB 361 RLHALGALLPGAGVOPGMHYSRLYRLAEELG 396

RESULT 6
 W97766
 ID W97766 standard; Protein: 396 AA.
 AC W97766;
 DT 21-MAY-1999 (first entry)
 DE Mouse Desert hedgehog (Dhh) protein.
 KW Desert hedgehog; Dhh protein; mouse; dopaminergic; GABA-nergic;
 KW ptc therapeutic; patched; signal transduction; Parkinson's disease;
 KW Huntington's disease; amyotrophic lateral sclerosis;
 KW cerebral ischemia; hypoxia; neuroprotective; therapy.
 OS Mus sp.
 PN W09904775-A2.
 PD 04-FEB-1999.
 PF 24-JUL-1998; U15419.
 PR 24-JUL-1997; US-900220.
 PA (ONTO-) ONTOGENY INC.
 PI Mahantappa NK, Miao N, Pang K, Wang M;
 DR WPI: 99-142578/12.
 DR N-PSDB: X072272.
 PT Increasing the survival of neuronal, dopaminergic and GABA-nergic
 PT cells - by using a ptc therapeutic such as a protein kinase
 PT inhibitor, or an agent derived from hedgehog polypeptides, useful in
 PT the treatment of Parkinson's disease
 PS Disclosure: Page 85-86; 138pp; English.
 CC This polypeptide is mouse Desert hedgehog protein (Dhh). The
 CC invention is based on the finding that hedgehog proteins are useful
 CC as protective agents in the treatment and prophylaxis of
 CC neurodegenerative disorders resulting from the loss of dopaminergic
 CC and/or GABA-nergic neurons, or the general loss of tissue from the
 CC substantia nigra. Exemplary disorders include Parkinson's disease,
 CC Huntington's disease (both claimed), amyotrophic lateral sclerosis
 CC and cerebral ischemia. The invention relates to hedgehog
 CC therapeutics (i.e. hedgehog polypeptides and gene therapy constructs
 CC 3-9. constructs encoding recombinant hedgehog polypeptides and
 CC trans-activation constructs for altering hedgehog gene regulatory
 CC sequences) and ptc therapeutics (i.e. agents which mimic the effect
 CC of naturally occurring hedgehog proteins on patched signaling)
 CC that are effective in both human and animal subjects. Human Ihh
 CC and Dhh polypeptides (see W97763-64) are preferred. The products
 CC can also be used for the maintenance of differentiated neurons in
 CC cultures, and to enhance the implantation of such neuronal cells in
 CC an animal. They can be used to prevent or treat neurodegenerative
 CC conditions arising from the use of certain drugs, and in the

CC prevention and/or treatment of hypoxia, e.g. as a neuroprotective
 CC agent.
 SQ Sequence 396 AA;
 Query Match 96.2%; Score 2008; DB 1; Length 396;
 Best Local Similarity 96.5%; Pred. No. 4.5e-208;
 Matches 382; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 MALLTNLPLCLALLALPAQSCGPGRGVRRRRARAKQVLPPLYKQVPGVPERTLGAS 60
 DB 1 MALPASLPLCLALLALSAQSCGPGRGVRRRRARAKQVLPPLYKQVPGVPERTLGAS 60
 QY 61 GAEGGVARGSERFRDLVNPNDIIFKDEENSGADRLMTERCKERVNALAIVANNMMPG 120
 DB 61 GAEGGVARGSERFRDLVNPNDIIFKDEENSGADRLMTERCKERVNALAIVANNMMPG 120
 QY 121 VALRTGEMDEGHHADSLHYEGRALDITTSDRRNKYGGLARLAVAGDFWYYESRN 180
 DB 121 VALRTGEMDEGHHADSLHYEGRALDITTSDRRNKYGGLARLAVAGDFWYYESRN 180
 QY 181 HHVSVKADNSLAVRAGCGCPGNATVRLMSGERKGLRELHGDWVLAADAGRVVPTVYL 240
 DB 181 HHVSVKADNSLAVRAGCGCPGNATVRLMSGERKGLRELHGDWVLAADAGRVVPTVYL 240
 QY 241 LFLDLDLORRASFVAVETERPRLKLLTPWHLVFARGPAPAGDFAVFARRLRAGDSV 300
 DB 241 LFLDLDLORRASFVAVETERPRLKLLTPWHLVFARGPAPAGDFAVFARRLRAGDSV 300
 QY 301 LAPGSDALPAPVARVAREEAVGFAPLTAHGTLLVNDVLASCVAVLSEHOMAHRAFAPL 360
 DB 301 LAPGSDALPAPVARVAREEAVGFAPLTAHGTLLVNDVLASCVAVLSEHOMAHRAFAPL 360
 QY 361 RLHALGALLPGAGVOPGMHYSRLYRLAEELG 396
 DB 361 RLHALGALLPGAGVOPGMHYSRLYRLAEELG 396

RESULT 7
 Y05511
 ID Y05511 standard; Protein: 396 AA.
 AC Y05511;
 DT 05-JUL-1999 (first entry)
 DE Mouse Desert hedgehog protein Dhh.
 KW Desert hedgehog; Dhh protein; mouse; hedgehog therapeutic;
 KW ptc therapeutic; patched; signal transduction; muscle atrophy;
 KW cachexia; muscular myopathy; myoblastic sarcoma; therapy.
 OS Mus sp.
 PN W09910004-A2.
 PD 04-MAR-1999.
 PF 28-AUG-1998; U17922.
 PR 29-AUG-1997; US-057394.
 PA (ONTO-) ONTOGENY INC.
 PI Bladgen CS, Currie PD, Hughes SM, Ingham PW;
 DR WPI: 99-243557/20.
 DR N-PSDB: X25099.
 PT A new method to regulate muscle growth
 PT Disclosure: Page 112-113; 130pp; English.
 PS The present sequence is mouse Desert hedgehog protein Dhh. The
 PS invention relates to a method for modulating the formation and/or
 PS maintenance of muscle tissue by ectopically contacting muscle
 PS cells, especially muscle stem/progenitor cells, in vitro or in
 PS vivo, with a hedgehog therapeutic (i.e. hedgehog polypeptides and
 PS gene therapy constructs) or ptc therapeutic (i.e. a small organic
 PS molecule that mimics the effect of hedgehog proteins on patched
 PS signaling, or activates or potentiates patched signaling) in an
 PS amount effective to alter the growth state of the treated cells.
 PS Also claimed is a method for treatment or prevention of disorders
 PS of, or surgical or cosmetic repair of, such muscle tissues, by
 PS administering a hedgehog polypeptide or ptc therapeutic. The
 PS disorder may be muscle atrophy, in particular skeletal muscle
 PS atrophy or cardiac muscle atrophy, cachexia, or muscular myopathy
 PS (all claimed). The hedgehog polypeptide or ptc therapeutic can

CC may be used as diagnostic agents for neuromuscular, autonomic or
 CC central nervous system disorders, and the gene may also be used in
 CC gene therapy. Antibodies generated from the protein may be used
 CC as therapeutic or research reagents.
 SQ Sequence 437 AA:

Query Match 58.3%; Score 1217; DB 1; Length 437;
 Best Local Similarity 58.7%; Pred. No. 9,66-123;
 Matches 249; Conservative 53; Mismatches 82; Indels 40; Gaps 12;

QY 7 LPLCLALLA-----LPAQSGPGRGPGVGRRRYARKOLVPLLYKQFVGPVPTLGASG 61
 DB 4 LLARCFVLTLASSLLVCPGLACGPGRG-FGKRHRPKK-LTPLAYKQFIRNVAKETLGASG 61
 QY 62 PAEGVARGSERFDLVPNYNDIIFKDENSGADRLMTTERCKERNALATAVMMMPGV 121
 DB 62 REGKITRNSERFKELTPNYNDIIFKDEENTGADRLMTORCKDKLNALATSVMMQPGV 121
 QY 122 RLRTVEGWDEDDHNSDLHYEGRALDITTSDDRNKYGLLARLAVEAGFDMVYVESRHH 181
 DB 122 RLRTVEGWDEDDHNSDLHYEGRALDITTSDDRNKYGLLARLAVEAGFDMVYVESRHH 181
 QY 182 VHVSKADNSLAVRAGGCPGNATVRLMSGERKGLRELHRGDMVLAADASGRVPTVLL 241
 DB 182 IHCVSKAENSVAAKSGCPGSAIVHLBEGGKTVLDLPGRDVLAAADOGRLTSDFLT 241
 QY 242 FLDRDLORRASFAVAVETMPPRKLLTPMHLVFAA---RGPARAGDFAPVFAARLRAG 297
 DB 242 FLDRDLORRASFAVAVETMPPRKLLTPMHLVFAA---RGPARAGDFAPVFAARLRAG 297
 QY 298 DSVLA---PGSD-ALRPARVARVA-REEAVGFAPLTAHGTLLVNDVLAACYAVLESROW 352
 DB 298 QRVYVAERGGDRLLPAVAHSVTLREEAGAYAPLTAHGTLLINRVLAACYAVLESROW 352
 QY 353 AHRAPAPRLALHA-LGALLP-----GGAU-----GPT-GMHWYSLTLVLA 391
 DB 353 AHRAPAPRLALHA-LGALLP-----GGAU-----GPT-GMHWYSLTLVLA 391
 QY 392 EELL 395
 DB 419 TWLL 422

RESULT 10
 W94471 standard; Protein: 437 AA.
 ID W94471 standard; Protein: 437 AA.
 AC W94471;
 DT 29-APR-1999 (first entry)
 DE Mouse Shh hedgehog protein sequence.
 KW Patched; hedgehog; ptc therapeutic; neuroprotective; neuronal cell;
 KW brain infarction; cerebral infarction; transient ischemic attack;
 KW stroke; cerebral infarct volume; spinal cord; oedema; trauma;
 KW haemorrhage; encephalomyelitis; coronary bypass; cerebral hypoxia.
 OS Mus sp.
 PN W09900117-A2.
 PD 26-JUN-1998; 013387;
 PF 26-JUN-1998; 013387;
 PR 27-JUN-1997; US-883656.
 PA (ONTO-) ONTOGENY INC.
 PI Mahanahappa NK;
 DR WPI: 99-095458/08.
 N-PSDB: X16185.
 PT Method for limiting damage to neurons caused by ischemic or epoxic
 PT conditions - is used for the treatment and prevention of e.g.
 PT cerebral infarction, stroke and transient ischemic attacks
 PS Disclosure: Page 68-70: 104pp. English.
 CC A method has been developed for limiting the damage to neuronal cells by
 CC ischemic or epoxic conditions by administering a ptc (patched)
 CC therapeutic agent to reduce cerebral infarct volume (CIV). Damage to
 CC neuronal cells can also be limited by administering a gene activation
 CC construct which recombines with the genomic hedgehog gene to provide a
 CC heterologous transcription regulator linked to the coding region of this

CC gene. Administration of the ptc therapeutic agent is used to protect
 CC cerebral tissues against ischemic injury; to treat cerebral infarct or
 CC ischemia, stroke (thrombotic or embolic) and transient ischemic
 CC attacks. It may also be used as a prophylactic in many other cases of
 CC injury to the brain or spinal cord, oedema caused by trauma, hemorrhage
 CC and encephalomyelitis, or in conjunction with (coronary bypass) surgery.
 CC Treatment (which may be prophylactic) is used where ischemic/epoxic
 CC conditions may cause cerebral hypoxia, or progressive loss of neurons
 CC due to oxygen depletion, including in patients with hypotension. The
 CC treatment reduces CIV by at least 25, particularly at least 70%. The
 CC present sequence represents a hedgehog sequence given in the present
 CC invention.
 SQ Sequence 437 AA:

Query Match 58.3%; Score 1217; DB 1; Length 437;
 Best Local Similarity 58.7%; Pred. No. 9,66-123;
 Matches 249; Conservative 53; Mismatches 82; Indels 40; Gaps 12;

QY 7 LPLCLALLA-----LPAQSGPGRGPGVGRRRYARKOLVPLLYKQFVGPVPTLGASG 61
 DB 4 LLARCFVLTLASSLLVCPGLACGPGRG-FGKRHRPKK-LTPLAYKQFIRNVAKETLGASG 61
 QY 62 PAEGVARGSERFDLVPNYNDIIFKDENSGADRLMTTERCKERNALATAVMMMPGV 121
 DB 62 REGKITRNSERFKELTPNYNDIIFKDEENTGADRLMTORCKDKLNALATSVMMQPGV 121
 QY 122 RLRTVEGWDEDDHNSDLHYEGRALDITTSDDRNKYGLLARLAVEAGFDMVYVESRHH 181
 DB 122 RLRTVEGWDEDDHNSDLHYEGRALDITTSDDRNKYGLLARLAVEAGFDMVYVESRHH 181
 QY 182 VHVSKADNSLAVRAGGCPGNATVRLMSGERKGLRELHRGDMVLAADASGRVPTVLL 241
 DB 182 IHCVSKAENSVAAKSGCPGSAIVHLBEGGKTVLDLPGRDVLAAADOGRLTSDFLT 241
 QY 242 FLDRDLORRASFAVAVETMPPRKLLTPMHLVFAA---RGPARAGDFAPVFAARLRAG 297
 DB 242 FLDRDLORRASFAVAVETMPPRKLLTPMHLVFAA---RGPARAGDFAPVFAARLRAG 297
 QY 298 DSVLA---PGSD-ALRPARVARVA-REEAVGFAPLTAHGTLLVNDVLAACYAVLESROW 352
 DB 298 QRVYVAERGGDRLLPAVAHSVTLREEAGAYAPLTAHGTLLINRVLAACYAVLESROW 352
 QY 353 AHRAPAPRLALHA-LGALLP-----GGAU-----GPT-GMHWYSLTLVLA 391
 DB 353 AHRAPAPRLALHA-LGALLP-----GGAU-----GPT-GMHWYSLTLVLA 391
 QY 392 EELL 395
 DB 419 TWLL 422

RESULT 11
 W97768 standard; Protein: 437 AA.
 ID W97768 standard; Protein: 437 AA.
 AC W97768;
 DT 21-MAY-1999 (first entry)
 DE Mouse Sonic hedgehog (Shh) protein.
 KW Sonic hedgehog; Shh protein; mouse; dopaminergic; GABA-ergic;
 KW ptc therapeutic; patched; signal transduction; Parkinson's disease;
 KW Huntington's disease; amyotrophic lateral sclerosis;
 KW cerebral ischemia; hypoxia; neuroprotective; therapy.
 OS Mus sp.
 PN W099004775-A2.
 PD 04-FEB-1999.
 PF 04-FEB-1999.
 PR 24-JUL-1997; US-900220.
 PA (ONTO-) ONTOGENY INC.
 PI Mahanahappa NK, Miao N, Pang K, Wang M;
 DR WPI: 99-142578/12.
 N-PSDB: X07274.
 PT Increasing the survival of neuronal, dopaminergic and GABA-ergic
 PT cells - by using a ptc therapeutic such as a protein kinase

PT inhibitor, or an agent derived from hedgehog polypeptides, useful in
PT the treatment of Parkinson's disease
PS Disclosure; page 89-91; 138pp; English.
CC This polypeptide is mouse Shh Sonic hedgehog protein. The
CC invention is based on the finding that hedgehog proteins are useful,
CC as protective agents in the treatment and prophylaxis of
CC neurodegenerative disorders resulting from the loss of dopaminergic
CC and/or GABA-ergic neurons, or the general loss of tissue from the
CC substantia nigra. Exemplary disorders include Parkinson's disease,
CC Huntington's disease (both claimed), amyotrophic lateral sclerosis
CC and cerebral ischaemia. The invention relates to hedgehog
CC therapeutics (i.e. hedgehog polypeptides and gene therapy constructs
CC e.g. constructs encoding recombinant hedgehog polypeptides and
CC trans-activation constructs for altering hedgehog gene regulatory
CC sequences) and ptc therapeutics (i.e. agents which mimic the effect
CC of naturally occurring hedgehog proteins on patched signalling)
CC that are effective in both human and animal subjects. Human inh
CC and dhb polypeptides (see W97763-64) are preferred. The products
CC can also be used for the maintenance of differentiated neurons in
CC cultures, and to enhance the implantation of such neuronal cells in
CC an animal. They can be used to prevent or treat neurodegenerative
CC conditions arising from the use of certain drugs, and in the
CC prevention and/or treatment of hypoxia, e.g. as a neuroprotective
CC agent.
SQ Sequence 437 AA:

Query Match 58.3%; Score 1217; DB 1; Length 437;
Best Local Similarity 58.7%; Pred. No. 9.6e-123;
Matches 249; Conservative 53; Mismatches 82; Indels 40; Gaps 12;

QY 7 LLPLCCALLA-----LPAQSGGPGRGVGRRRYARKOLVPLLYKQFVPGVPTIGASG 61
DB 4 LLARCFVLTLASSLLVCGLACGPGRG-FGKRRHPKK-LTPLATKQFIPNVAEKTIGASG 61

QY 62 PAEGVAVGSGRFRDLVYNNPDIIFKDENSGADRLMTERCKEYVNALATAVNMMPGV 121
DB 62 RYEGKITNSRPFELTNNYNDIIFKDENGADRLMTQCKXKLNALATAVNMMPGV 121

QY 122 RLRYTEGWDGHHADSLHREGRALDITTSRDNRKYGTLARLAVEAGFDMVYYESRNH 181
DB 122 RLRYTEGWDGHHSESLHREGRAVDITTSRDNRKYGTLARLAVEAGFDMVYYESRNH 181

QY 182 VHVSVKADNSLAVRAGGCGFNATVRLMSGKRGRLRELRGDVLAADASGRVPTVLL 241
DB 182 IHCYSKAEVNSVAAGSGCGFPGSATVHLEGGGKLVKDLRPDRVLAADQGRLLYSDFLT 241

QY 242 FLDRDLQRRASVAVETEMPRKLLTPMHLVFAA-----RGAPAPGDFAVFARLRAG 297
DB 242 FLDRDEGAKKVFYVETLEPRERLLTLAHLFLFAPHNDSGTGP-----SALFASRVPRG 298

QY 298 DSVLA---PGGD-ALRPARVARVA-REBAVGFAPLTAHGTLLVNDVLAACYAVLESHQW 352
DB 298 QRVYVAERGGDRLLPAAVHSVTLREBEAGAYAPLTAHGTLLIRVLAACYAVLESHQW 358

QY 353 AHRAPAPRLRLHA-LGALLP-----GAV-----QPT-GMHVYSRLLYRLA 391
DB 353 AHRAPAPRLRLHA-LGALLP-----GAV-----QPT-GMHVYSRLLYRLA 391

QY 392 EELL 395
DB 419 TWLL 422

RESULT 12
Y05513
ID Y05513 standard; Protein; 437 AA.
AC Y05513;
DT 05-JUL-1999 (first entry)
DE Mouse Sonic hedgehog protein Shh.
KW Sonic hedgehog; Shh protein; mouse; hedgehog therapeutic;
ptc therapeutic; patched; signal transduction; muscle atrophy;
cachexia; muscular myopathy; myoblastic sarcoma; therapy.

OS Mus sp.
PN WO991004-A2.
PD 04-MAR-1999.
PF 28-AUG-1998; U17922.
PR 29-AUG-1997; US-057394.
PA (ONTOC) ONTOCENT INC.
PI Bladgen CS, Currie PD, Hughes SM, Ingham PW;
DR WPI: 99-243557/20.
DR N-PSDB: X25101.
PT A new method to regulate muscle growth.
PS Disclosure; page 115-116; 130pp; English.
CC The present sequence is mouse Sonic hedgehog protein Shh. The
CC invention relates to a method for modulating the formation and/or
CC maintenance of muscle tissue by ectopically contacting muscle
CC cells, especially muscle stem/progenitor cells, in vitro or in
CC vivo, with a hedgehog therapeutic (i.e. hedgehog polypeptides and
CC gene therapy constructs) or ptc therapeutic (i.e. a small organic
CC molecule that mimics the effect of hedgehog proteins on patched
CC signalling, or activates or potentiates patched signalling) in an
CC amount effective to alter the growth state of the treated cells.
CC Also claimed is a method for treatment or prevention of disorders
CC of, or surgical or cosmetic repair of, such muscle tissues, by
CC administering a hedgehog polypeptide or ptc therapeutic. The
CC disorder may be muscle atrophy, in particular skeletal muscle
CC atrophy or cardiac muscle atrophy, cachexia, or muscular myopathy
CC (all claimed). The hedgehog polypeptide or ptc therapeutic can
CC inhibit growth of myoblastic-derived tissue to provide treatment of
CC hyperplastic or neoplastic growth of muscle tissue such as in
CC myoblastic sarcoma (also claimed). The hedgehog therapeutic
CC preferably comprises at least a bioactive extracellular portion of
CC a hedgehog protein (see Y05510-19) encoded by a vertebrate hedgehog
CC gene (see X25098-107), especially a human hedgehog gene.
SQ Sequence 437 AA:

Query Match 58.3%; Score 1217; DB 1; Length 437;
Best Local Similarity 58.7%; Pred. No. 9.6e-123;
Matches 249; Conservative 53; Mismatches 82; Indels 40; Gaps 12;

QY 7 LLPLCCALLA-----LPAQSGGPGRGVGRRRYARKOLVPLLYKQFVPGVPTIGASG 61
DB 4 LLARCFVLTLASSLLVCGLACGPGRG-FGKRRHPKK-LTPLATKQFIPNVAEKTIGASG 61

QY 62 PAEGVAVGSGRFRDLVYNNPDIIFKDENSGADRLMTERCKEYVNALATAVNMMPGV 121
DB 62 RYEGKITNSRPFELTNNYNDIIFKDENGADRLMTQCKXKLNALATAVNMMPGV 121

QY 122 RLRYTEGWDGHHADSLHREGRALDITTSRDNRKYGTLARLAVEAGFDMVYYESRNH 181
DB 122 RLRYTEGWDGHHSESLHREGRAVDITTSRDNRKYGTLARLAVEAGFDMVYYESRNH 181

QY 182 VHVSVKADNSLAVRAGGCGFNATVRLMSGKRGRLRELRGDVLAADASGRVPTVLL 241
DB 182 IHCYSKAEVNSVAAGSGCGFPGSATVHLEGGGKLVKDLRPDRVLAADQGRLLYSDFLT 241

QY 242 FLDRDLQRRASVAVETEMPRKLLTPMHLVFAA-----RGAPAPGDFAVFARLRAG 297
DB 242 FLDRDEGAKKVFYVETLEPRERLLTLAHLFLFAPHNDSGTGP-----SALFASRVPRG 298

QY 298 DSVLA---PGGD-ALRPARVARVA-REBAVGFAPLTAHGTLLVNDVLAACYAVLESHQW 352
DB 298 QRVYVAERGGDRLLPAAVHSVTLREBEAGAYAPLTAHGTLLIRVLAACYAVLESHQW 358

QY 353 AHRAPAPRLRLHA-LGALLP-----GAV-----QPT-GMHVYSRLLYRLA 391
DB 353 AHRAPAPRLRLHA-LGALLP-----GAV-----QPT-GMHVYSRLLYRLA 391

QY 392 EELL 395
DB 419 TWLL 422

RESULT 13

W61488
ID W61488 standard; Protein: 437 AA.
AC W61488;
DT 20-OCT-1998 (first entry)
DE Mouse sonic hedgehog (shh) protein.
KW Hedgehog polypeptide; neuronal cell proliferation; zebrafish; shh;
KW cholesterol biosynthesis; pituitary gland gene expression; mouse;
KW liggy-winkle hedgehog protein; twm; sonic hedgehog; major limb trauma;
KW neuronal degeneration; nerve-sparing agent; Smith-Lemli-Opitz syndrome.
OS Mus sp.
PN W09830576-A1.
PT 16-JUL-1998.
PF 07-OCT-1997; U15753.
PR 02-OCT-1997; U5-061323.
PA 07-OCT-1996; US-729743.
PI (UYUO) UNIV JOHNS HOPKINS SCHOOL OF MEDICINE.
PI Beachy PA, Porter JA;
DR WPI: 98-399053/34.
PT New hedgehog-derived poly(peptide)s - used to develop products for
modulating proliferation or differentiation of neuronal cells,
PT cholesterol biosynthesis or transport or expression of pituitary
PT gland gene(s)
PS Claim 13: Pages 153-154; 210pp; English.
CC This represents a mouse sonic hedgehog (shh) protein sequence. The
CC invention provides methods and compounds for modulating proliferation or
CC differentiation of neuronal cells, cholesterol biosynthesis or transport
CC or expression of pituitary gland genes. The method for affecting
CC cholesterol biosynthesis or transport in a cell comprises contacting a
CC cell with a compound that affects hedgehog, thereby affecting cholesterol
CC biosynthesis or transport. The methods for inhibiting the neural inducing
CC activity of a hedgehog polypeptide in cells, and for inducing pituitary
CC gland gene expression utilizes sequences selected from a zebrafish
CC liggy-winkle hedgehog (twm) protein (W61487) or a mouse shh
CC (shh) protein (W61488), a chicken shh protein (W61487) or a mouse shh
CC protein (W61488). The products and methods provide for compounds which
CC can affect hedgehog activity. They can be used for treating disorders
CC which arise from neuronal degeneration or abnormal function. They can
CC also be used as nerve-sparing agents or in restoring or promoting
CC appropriate patterning during the healing of major limb trauma. They can
CC also be used for treating Smith-Lemli-Opitz syndrome. The products can
CC also be used for detection and diagnosis.
SO Sequence 437 AA:

Query Match 58.1%; Score 1214; DB 1; Length 437;
Best Local Similarity 58.5%; Pred. No. 2e-122;
Matches 248; Conservative 54; Mismatches 82; Indels 40; Gaps 12;

QY 7 LPLPCCLALLA-----LPAQSGGPGRGVGRRRYARKQVLPVLYKQFVGVPEBRTGASG 61
DB 4 LLAARFVLITASSLLVCPGLACGPGRG-FGKRHRPKR-LTPPLAYKQFIPVNAEKTUGASG 61
QY 62 PAEGVARSSEFFRDLPVNPDIIFKDEENSGADRLTECKERKERNALAIYVMMWPGV 121
DB 62 REGKRTNRSERFELTPYNPDIIFKDEENSGADRLTECKERKERNALAIYVMMWPGV 121
QY 122 RLRTVEGMEDEGHNAODSLHYEGRALDITTSDRNRKYYGLARLAVAGFDWYVESRNH 181
DB 122 KLRTEGMEDEGHNSHESLHYEGRAVDITTSDRNRKYYGLARLAVAGFDWYVESRNH 181
QY 182 VAVSYKADNSLAVRAGGCPFGNATVRLMSGERKGLRELHSGDMVTLAADSAGRVVPT 241
DB 182 IHCSTVAENSVAKSGGCPFGSATVRLHEOGGKLVKDLRPGRVTLAADOGRLYSDFLT 241
QY 242 FLDRDQKARASFAVATEMPRRKLLTPMHLVFAA-----RGPAPAGDPAFARLRMG 297
DB 242 FLDRDQKARVYVITTELRERKLLTAHLFLVAPHNDGPPGP--SALFASVVRG 298
QY 298 DSVLA---DGGD-ALRPAVARVA--REEAVGYEARPLTAHGTLLVNDVLAACYAVLESHQW 352
DB 299 GRVYVAERGGDRRLRPAVHVSYTLREEDAGAVAPLTAAGTLLINNVLASCAVIEHSHW 358
QY 353 AHRAPRRLRLHA-LGALLP-----GGAV-----OPT-GMHWYSRLLYRLA 391

DB 359 AHRAPRRLRLHAALAPARTDGGGSGIPAAQASATANGAEPTAGIHWYSQLLYHIG 418
QY 392 EELL 395
DB 419 TWLL 422

RESULT 14
R77338
ID R77338 standard; Protein: 425 AA.
AC R77338;
DT 22-FEB-1996 (first entry)
DE Chicken sonic hedgehog protein.
KW Chicken; sonic hedgehog protein; stage 22/22 limb bud; probe;
KW primer; diagnostic; nervous system disorder; gene therapy;
KW antibody.
OS Gallus domesticus.
PN Key
PT Location/Qualifiers
FT Peptide 1..26
FT /note= "putative signal peptide"
FT peptide 27..32
FT /note= "conserved sequence (R77349)"
FT modified site 282
FT /note= "N-linked glycosylation site"
FT W09518856-A1.
PN 13-JUL-1995.
PD 30-DEC-1994; U14992.
PR 30-DEC-1993; US-176427.
PR 14-DEC-1994; US-356060.
PA (HARD) HARVARD COLLEGE.
PI (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
PI Ingham PW, McMahon AP, Tablin CJ;
DR WPI: 95-255060/33.
DR N-PSDB: Q91636.
PT Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful
PT to treat degenerative nervous system disorder(s) and in gene
PT therapy.
PS Claim 17: Page 133-35; 210pp; English.
CC The sequence represents a chicken sonic hedgehog protein.
CC homologous to a Drosophila hedgehog protein (R77337), and is
CC encoded by a cDNA isolated from a stage 22/22 limb bud cDNA
CC library. Probes and primers derived from the sonic hedgehog gene
CC may be used as diagnostic agents for neuromuscular, autonomic or
CC central nervous system disorders, and the gene may also be used in
CC gene therapy. Antibodies generated from the protein may be used
CC as therapeutic or research reagents.
SO Sequence 425 AA:

Query Match 57.4%; Score 1198; DB 1; Length 425;
Best Local Similarity 59.4%; Pred. No. 1e-120;
Matches 244; Conservative 56; Mismatches 91; Indels 20; Gaps 10;

QY 1 MALLTNLP---LCLALLAPAOQSGGPGRGVGRRRYARKQVLPVLYKQFVGVPEBRTL 57
DB 4 MLTRILRLVGRFC--ALLVSSGLTCGPGRG-IGKRHRPKR-LTPPLAYKQFIPVNAEKT 59
QY 58 GASGPAEGRVARGSRFDLPVNPDIIFKDEENSGADRLTECKERKERNALAIYVMMW 117
DB 60 GASGPEKRTNRSERFELTPYNPDIIFKDEENSGADRLTECKERKERNALAIYVMMW 119
QY 118 WPGVRLRTVEGMEDEGHNAODSLHYEGRALDITTSDRNRKYYGLARLAVAGFDWYVE 177
DB 120 WPGVRLRTVEGMEDEGHNSHESLHYEGRAVDITTSDRNRKYYGLARLAVAGFDWYVE 179
QY 178 SRNHVAVSKADNSLAVRAGGCPFGNATVRLMSGERKGLRELHSGDMVTLAADSAGRVPT 237
DB 180 SKAHTHCSVKAENSVAKSGGCPFGSATVRLHEHGTGLVKDLRPGDRLAADAGRLYLS 239
QY 238 PVLLFLDRDQKARASFAVATEMPRRKLLTPMHLVFAA--RGPAPAGDPA--PVARR 293
DB 240 DFLTFDRDMSRKLFLYVETROPARLLTAHLFLVAPHOHNOSEATGTSGOALFASN 299

QY 294 LRAQDS--VIAPGGDALPAPARVARVA-REAVGVFAPLTAHGTLLVNDVASCYAVIESH 350
 DB 300 VKPGORVYVGEQGOQLLPASVSHSVSLREASGAYAPLTAOGTLLINRYLASCYAVIEEH 359
 QY 351 QMAHRAFAPLRLHA-LGALLPGAV-----OPTGMHWSRLTYRLAEELL 395
 DB 360 SMAHMAFAPRLAOGLLAALCPDGAIPTAATTGTHWISRLTYRIGSWYL 410

RESULT 15

W94468
 ID W94468 standard; Protein; 425 AA.
 AC W94468;
 DT 29-APR-1999 (first entry)
 DE Chicken Shh hedgehog protein sequence.
 KM Patched; hedgehog; ptc therapeutic; neuroprotective; neuronal cell;
 KM brain infarction; cerebral infarction; transient ischaemic attack;
 KM stroke; cerebral infarct volume; spinal cord; oedema; trauma;
 KM haemorrhage; encephalomyelitis; coronary bypass; cerebral hypoxia.
 OS Gallus sp.
 PN MO9900117-A2.
 PD 07-JAN-1999.
 PF 26-JUN-1998; U13387.
 PR 27-JUN-1997; US-883656.
 PA (ONTO-) ONTOGENY INC.
 PI Mahantappa NK;
 DR WPI: 99-095458/08.
 DR N-PSDB: X16182.
 PT Method for limiting damage to neurons caused by ischaemic or epoxic
 PT conditions - is used for the treatment and prevention of e.g.
 PT cerebral infarction, stroke and transient ischaemic attacks
 PS Disclosure; Page 63-64; 104pp; English.
 CC A method has been developed for limiting the damage to neuronal cells by
 CC ischaemic or epoxic conditions by administering a ptc (patched)
 CC therapeutic agent to reduce cerebral infarct volume (CIV). Damage to
 CC neuronal cells can also be limited by administering a gene activation
 CC construct which recombines with the genomic hedgehog gene to provide a
 CC heterologous transcription regulator linked to the coding region of this
 CC gene. Administration of the ptc therapeutic agent is used to protect
 CC cerebral tissues against ischaemic injury; to treat cerebral infarct or
 CC ischaemia, stroke (thrombotic or embolic) and transient ischaemic
 CC attacks. It may also be used as a prophylactic in many other cases of
 CC injury to the brain or spinal cord, oedema caused by trauma, haemorrhage
 CC and encephalomyelitis, or in conjunction with (coronary bypass) surgery.
 CC Treatment (which may be prophylactic) is used where ischaemic/epoxic
 CC conditions may cause cerebral hypoxia, or progressive loss of neurons
 CC due to oxygen depletion, including in patients with hypotension. The
 CC treatment reduces CIV by at least 25, particularly at least 70%. The
 CC present sequence represents a hedgehog sequence given in the present
 CC invention.
 SO Sequence 425 AA;

Query Match 57.4%; Score 1198; DB 1; Length 425;

Best Local Similarity 59.4%; Pred. No. 1e-120;

Matches 244; Conservative 56; Mismatches 91; Indels 20; Gaps 10;

QY 1 MALLTNLTP---LCCIALALPAPSCGPGRGVARRARAKQVPLLYKQVPGVPEKTL 57
 DB 4 MLLTLRLVGFIC--ALVSSGLTCGPRG-IGKRHRPK-LTPPLAYKQFIPNVAEKTLL 59
 QY 58 GASGPAEGRVARGSERFDLYPNTNPDIIFKDEENSGADRLMTERCKREYNALAIIVNM 117
 DB 60 GASGRYBEKTRNSERFELTPNPNPDITTFKDEENTGADRLMTQCKDKLNALISVNMQ 119
 QY 118 WPGVRLRVTEGMDGDGHAAQDSLHYGRALDITTSDRDRNKRYGLRLAVEAGFDWVY 177
 DB 120 WPGVRLRVTEGMDGDGHSESLHYEGRAVDITTSDRDRSKYGMRLAVEAGFDWVY 179
 QY 178 SRNHVHSVKAADNSLAVAGCGFPGNATVRLMSGERKGLRELHAGDWYLAADASGRVPT 237
 DB 180 SKAHHCYSKAVENSVAAKSGCGFPGSATVHLHGQTKLVKDLSPGDRVLAADADGRLLYS 239

QY 238 PVLLFDRLQORASVAVETEMPFRKLLTPMHLVFAA--RGPAPAGDFA--PVFARR 293
 DB 240 DFLTFDRMDSKRLFYVETRQPARRLTLTAHLLFVAPQHNSATGTSQALFASN 299
 QY 294 LRAQDS--VIAPGGDALPAPARVARVA-REAVGVFAPLTAHGTLLVNDVASCYAVIESH 350
 DB 300 VKPGORVYVGEQGOQLLPASVSHSVSLREASGAYAPLTAOGTLLINRYLASCYAVIEEH 359
 QY 351 QMAHRAFAPLRLHA-LGALLPGAV-----OPTGMHWSRLTYRLAEELL 395
 DB 360 SMAHMAFAPRLAOGLLAALCPDGAIPTAATTGTHWISRLTYRIGSWYL 410

Search completed: June 5, 2000, 08:16:20
 Job time: 2585 sec

Thu Jun 8 15:52:54 2000

us-08-900-220-17.rag

Page 10

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2000, 08:16:26 ; Search time 26.41 Seconds

(without alignments)
1078.998 Million cell updates/sec

Title: US-08-900-220-16

Sequence: 1 MSPARLRPRHPCVLVLLIL.....GRLLEGGSFHPLGSGAGS 411

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 225878 seqs, 69334122 residues

Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1539	70.8	406	13	057567 notophthalm
2	1285	59.1	434	13	057404 pleurodeles
3	1253.5	58.1	414	13	09w709 parallelthy
4	1086.5	50.0	442	13	073803 fugu rubrip
5	1076.5	49.5	415	5	017499 branchiosto
6	914	42.0	410	5	061676 lytechinus
7	864	41.0	139	11	09w229 rattus nov
8	701	32.2	139	6	09xsi6 bos taurus
9	681	31.3	138	13	09w6c1 eleutheroda
10	665	30.6	150	13	09ygv7 ambystoma m
11	597	27.4	185	5	096699 junonia coe
12	510.5	23.5	129	11	09wmp6 rattus nov
13	472	21.7	119	13	042128 oryzias lat
14	397	18.3	88	13	09ygu3 brachydanio
15	389	17.9	80	13	042441 oryzias lat
16	239	11.0	137	13	042234 coturnix co
17	174.5	8.0	615	5	091573 caenorhabdi
18	173.5	8.0	868	5	09xvi4 caenorhabdi
19	170.5	7.8	1021	5	09xuv2 caenorhabdi
20	161	7.4	790	5	022872 caenorhabdi

21	161	7.4	1226	5	021835 caenorhabdi
22	154.5	7.1	1207	5	021535 caenorhabdi
23	148.5	6.8	481	5	045992 caenorhabdi
24	148	6.8	485	5	094128 caenorhabdi
25	140	6.4	557	5	094129 caenorhabdi
26	138.5	6.4	550	5	094130 caenorhabdi
27	138.5	6.4	629	5	045273 caenorhabdi
28	123.5	5.7	205	5	023193 caenorhabdi
29	105	4.8	1997	2	005647 streptomyce
30	102	4.7	751	5	021754 caenorhabdi
31	100.5	4.6	478	5	043955 leishmania
32	99	4.6	1013	2	053499 mycobacteri
33	98.5	4.5	316	2	0928v4 streptomyce
34	98	4.5	1222	1	09yc75 aeropyrium p
35	97	4.5	1551	4	014160 homo sapien
36	97	4.5	7463	2	0924x6 streptomyce
37	96.5	4.4	54	13	042233 coturnix co
38	96.5	4.4	647	12	039278 equine hept
39	96	4.4	8563	2	054297 streptomyce
40	95.5	4.4	299	4	09y6k2 homo sapien
41	95.5	4.4	846	2	053152 mycobacteri
42	95	4.4	950	2	050470 mycobacteri
43	95	4.4	1616	2	096285 mycobacteri
44	95	4.4	2117	2	092429 streptomyce
45	94.5	4.3	257	4	09y2u4 homo sapien

ALIGNMENTS

RESULT	1	PRELIMINARY;	PRT;	406 AA.
ID	057567			
AC	057567;			
DT	01-JUN-1998 (TEMBLrel. 06, Created)			
DT	01-JUN-1998 (TEMBLrel. 06, Last sequence update)			
DT	01-NOV-1999 (TEMBLrel. 12, Last annotation update)			
DE	HEDEHOG SEGMENT POLARITY HOMOLOG.			
OS	Notophthalmus viridescens (Eastern newt) (Triturus viridescens).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;			
OC	Batrachia; Caudata; Salamandridae; Salamandridae; Notophthalmus.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	STARK D.R., GATES P.B., BROCKES J.P., FERRETTI P.;			
RL	Dev. Dyn. 0:0-0(1998).			
DR	EMBL: AF047466; AAC03108.1; -			
DR	HSSP: Q62226; LVHH.			
DR	PFAM: PF01085; HH_signal; 1.			
DR	PFAM: PF01079; HntL; 1.			
DR	PRINTS: PR00632; SONICHOG.			
DR	SEQUENCE 406 AA; 45072 MW; 9D0FFA76 CRC32;			
SO				
Query Match	70.8%;	Score 1539;	DB 13;	Length 406;
Best Local Similarity	74.1%;	Pred. No. 2.2e-117;		
Matches 292;	Conservative 31;	Mismatches 69;	Indels 2;	Gaps 2;
Db	18 LLLVPAAMGGPGRVVGRRRRPRKTVPLAYKQSPNVPEKTLGASGRGKTAASER	77		
Db	15 LLLGGPGLGCGPGRVIGRRPRPR-LIPUSYKFLHVEPKTLGASGRGKTAASER	73		
QY	78 FKELTLPVNPDIIFKDEENTGADRLMTQRCRDLNSLAISVMNQPVKRLVTEGMEDG	137		
Db	74 FKELTLPVNPDIIFKDEENTGADRLMTQRCRDLNSLAISVMNQPVKRLVTEGMEDG	133		
QY	138 HHSSESLLHSEGRAVDITTSRDRNKYGLARLAVAGFDWYVESKAHVHCSVSEHSA	197		
Db	134 HHSDESLLHSEGRAVDITTSRDRNKYGLARLAVAGFDWYVESKAHVHCSVSEHSA	193		
QY	198 AKTGCCPPARALATLESGEKIPADIPEGRHVLCDMDGGRGRTYSDFILFDROSTAVKE	257		
Db	194 AKTGCCPPARALATLESGEKIPADIPEGRHVLCDMDGGRGRTYSDFILFDROSTAVKE	253		


```

QY 258 QVETODPPRRRLATPAHLFTADNHTEPARFRATFASHVOPGOYVYAVGPLOPARV 317
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 254 YVETDPPRRRLATPAHLFTADNHTVPDTBSTVFAHVGOYILLEGVLGLOPARV 313
QY 318 AAVSTHVALGAVAPLTKHGTLVVEDVASCFAAVADHHLAQLAFWPLRLFHSLANGSWTP 377
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 314 VSTTGTDSGAVAPLTKHGTLVDDVYVSCFANVQKHQALAFWPLRLKHSVGRETOP 373
QY 378 GEGVHWYPOLLXRLGRLLLEEGSFHPLGMSGAGS 411
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 374 -EGMHWYSLLXRLGRLLLEEGSFHPLGMSGAGS 406

```

```

RESULT 2
ID 057404 PRELIMINARY: PRT: 434 AA.
AC 057404;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE SONIC HEDGEHOG-RELATED PROTEIN.
GN PW-SHH.
OS Pleurodeles waltlilii (Iberian ribbed newt).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC Batrachia; Caudata; Salamandroides; Salamandridae; Pleurodeles.
RN [1]
RP SEQUENCE FROM N.A.
RA CAUDIT X., NICOLAS S., LE PARCO Y.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF003532; AAB94412.1;
DR HSSP: 062236; 1VHH.
DR PFAM: PF01085; HH_Signal; 1.
DR PFAM: PF01079; Hnt; 1.
DR PRINTS: PRO0632; SONICHHOG.
SQ SEQUENCE 434 AA; 48421 MW; AFEDDF3 CRC32;

```

Query Match 59.1%; Score 1285; DB 13; Length 434;
 Best Local Similarity 61.2%; Pred. No. 1.1e-96;
 Matches 255; Conservative 46; Mismatches 90; Indels 26; Gaps 5;

```

QY 20 LVYPAAMGCGPGRVGSRRRPRKLVPLAYKQSPNVPEKTGASRGYEGKIARSSRRK 79
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 19 LVYPLGCGPGRGIGORRR-PQKTLPLAYKQIPNVSEKTLGASRGYEGKITRNSERQ 77
QY 80 ELTPNPNPDIIFKDEENTGADRLMTQRCRDLNSLAISVNMQMPGVKRLVTEGMDGHH 139
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 78 ELTPNPNPDIIFKVEENTGADRLMTQRCRDLNSLAISVNMQMPGVKRLVTEGMDGHH 137
QY 140 SEESLHYEGRAVDITTSDDRNKYGILLARLAVAGFDWYTESKAHVCSYKSEHSAAK 199
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 138 SEESLHYEGRAVDITTSDDRNKYGILLARLAVAGFDWYTESKAHVCSYKSEHSAAK 197
QY 200 TGSCFPAGAOVRLSGARVALSAVRPGDRLVLAGEDGSPFSDVLIJLDEPRRLAFOV 259
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 198 SGCGCFGSATVTEGQVRIYVKGKGNRYLADVDEGRILYSPFLFMDDEETAKRVFY 257
QY 260 IETQDPPRRRLATPAHLFTADNHTEPAA-----RFRATFASHVOPGOYVYAVG 308
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 258 IETSLPRERLRTAHLLEFAQEHFPGNASAGNFRSKFGHRFRSMPASSVRPHVLTED 317
QY 309 VP--GIQPARVAVSTHVALGAPLTKHGTLVVEDVASCFAAVADHHLAQLAFWPLRL 366
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 318 REGRGREATVDVYLEEGAVGAPVTAHGTVIDRYLASCYAVIEHSEVAHMAAFAPLRY 377
QY 367 -----FHSLANGSWTP-----GEGVHWYPOLLXRLGRLLLEEGSFHPLGMSGAGS 411
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 378 GEGILSFSPQDISSHSPPAPSOAEGVHWYSLLXRLGRLLLEEGSFHPLGMSGAGS 434

```

RESULT 3
 Q9W709 PRELIMINARY: PRT: 414 AA.
 ID Q9W709

```

AC Q9W709;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE SONIC HEDGEHOG.
GN SHH.
OS Paraliichthys olivaceus (Flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
OC Pleuronectiformes; Pleuronectoidel; Bothidae; Paraliichthys.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 99238226.
RA SUZUKI T., ICHIRO O., KUROKAWA T.;
RT "Retinoic acid given at late embryonic stage depresses sonic hedgehog
RT and Hoxd-4 expression in the pharyngeal area and induces skeletal
RT malformation in flounder (Paraliichthys olivaceus) embryos."
RL Dev. Growth Differ. 41:143-152(1999).
DR EMBL: AB029748; BAA82360.1;
SQ SEQUENCE 414 AA; 45945 MW; ELFB12EE CRC32;

```

Query Match 58.1%; Score 1263.5; DB 13; Length 414;
 Best Local Similarity 62.6%; Pred. No. 5.6e-95;
 Matches 251; Conservative 47; Mismatches 92; Indels 11; Gaps 6;

```

QY 16 LLLLVPAMGCGPGRVGSRRRPRKLVPLAYKQSPNVPEKTGASRGYEGKIARSS 75
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 12 VCLSLVSSGCGGPGRGYG-RRHPKTLPLAYKQIPNVSEKTLGASRGYEGKITRNS 70
QY 76 ERFKELTPYNDIIFKDEENTGADRLMTQRCRDLNSLAISVNMQMPGVKRLVTEGMD 135
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 71 ERFKELTPYNDIIFKDEENTGADRLMTQRCRDLNSLAISVNMQMPGVKRLVTEGMD 130
QY 136 DGHSEESIHYGRAVDITTSDDRNKYGILLARLAVAGFDWYTESKAHVCSYKSEHS 195
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 131 DGHSEESIHYGRAVDITTSDDRNKYGILLARLAVAGFDWYTESKAHVCSYKSEHS 190
QY 196 AAATGCGCFPAGAOVRLSGARVALSAVRPGDRLVLAGEDGSPFSDVLIJLDEPRRLR 255
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 191 VAAKSGGCGPGSSVTYLDQGTCKPKVALQTDGRVLAADHAGQPVYTDITMFDQOSTRR 250
QY 256 ACQVETODPPRRRLATPAHLFTA-DNHTEPAR-FRATFASHVOPGOYVYAVGPLO 313
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 251 LEVYLET-DSGKITLTPAHLFTVGHNSSTERNHGMKSAVFSAQVPOQYFVLDIERQ 309
QY 314 PARVAVSTHVALGAVAPLTKHGTLVVEDVASCFAAVADHHLAQLAFWPLRLFH---SL 370
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 310 PVTAKRIYQHEGSAFAPVTAAGTYVVDQVYLASCYAVIODEHLAALAPVRLAHVWSL 369
QY 371 AWGSMTPG---EGVHWYPOLLXRLGRLLLEEGSFHPLGMSGAGS 407
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 370 LFSQPOASAQKQGVHWYSKILQLGTWLDHSHIPLGMS 410

```

```

RESULT 4
ID 073803 PRELIMINARY: PRT: 442 AA.
AC 073803;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE FUGU HEDGEHOG.
GN FHR.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
OC Tetraodontiformes; Tetraodontidae; Tetraodontidae; Fugu.
RN [1]
RP SEQUENCE FROM N.A.
RA GELNER K., BRENNER S.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF056116; AAC34384.1;

```



```

OY 310 PGLQAPARVAVSTHVALGVAAPLTKHGTLVYEDVYASCRAAAVDHDLAQLAPLPLF--- 368
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 308 RGVARMKAVSVYTRIGHTRVAPVPTKQSGSYLVIGVAVSYVMKDEWVAHASFAPIMRYTY 367
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 368 --HSLAWGSMTPG--EGVHWYPOLLYRLGLRLLEEGSF 401
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 368 ISHMGITDITDTCGEQRYHWYTGGLTKLKYMSDLKF 405
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 7
O9WV29
ID O9WV29 PRELIMINARY; PRT: 177 AA.
AC O9WV29.
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE INDIAN HEDGEHOG PROTEIN (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
[1]
RN RN RP
SEQUENCE FROM N.A.
RA STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY.
RC GARGES P.L., MEYER R.A. JR., BROWN C.A., PRICE D.K.;
RT "Indian hedgehog in rat.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF162914; AAC45372.1; -.
FT NON_TER 1
FT NON_TER 177
FT NON_TER 177
SEQUENCE 177 AA; 19739 MW; 24CFI044 CRC32;

```

Query Match	41.6%	Score 904	DB 11	length 177
Best Local Similarity	96.6%	Pred. No. 2.9e-66		
Matches 171: Conservative	4	Mismatches 2	Indels 0	Gaps 0
QY	102	LMTQCRKRLNSLAIASVNMQMPGVKLRYTEGWDENDEGHSEESLHYEGRVADITTSDDRBN	161	
Db	1	LMTQCRKRLNSLAIASVNMQMPGVKLRYTEGWDENDEGHSEESLHYEGRVADITTSDDRBN	60	
QY	162	KYGLIARLAVEAGFDWVYYESKAHVCSVKSESHAARKTGCFPPAGQVRLSEGARVALS	221	
Db	61	KYGLIARLAVEAGFDWVYYESKAHVCSVKSESHAARKTGCFPPAGQVRLSEGARVALS	120	
QY	222	AVRPEDRYLAVGEGDSPTFSVDYLITLDRPHRLARFOYIETODPPRRLLTPAHLLF	278	
Db	121	AVRPEDRYLAVGEGDNPFTFSVDYLITLDRPHRLARFOYIETODPPRRLLTPAHLLF	177	
RESULT	8			
Q9XS16	PRELIMINARY	PRT	139 AA.	
AC	Q9XS16			
DT	01-NOV-1999 (TREMBLREL. 12, Created)			
DT	01-NOV-1999 (TREMBLREL. 12, Last sequence update)			
DT	01-NOV-1999 (TREMBLREL. 12, Last annotation update)			
DE	SONIC HEDGEHOG (FRAGMENT).			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
OC	Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae;			
OC	Bovinae; Bos.			
NN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=TOOTH GERM;			
RA	KOYAMA E., IWAMOTO M., OHMORI T., KURISU K., WU C., OOKURA T.,			
RA	BASHIR M.M., TUCKER T., PACIFICI M.;			
RT	"Development of Stratum intermedium and its role as a Sonic Hedgehog-			
RT	signaling structure during odontogenesis."			
RL	Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.			
FR	EMBL: AF144100; AAD33926.1; -			
FT	NON_TER	1	1	
FT	NON_TER	139	139	

50	SEQUENCE	139	AA;	15961	MM;	32D3F025	CRC32;
	Query Match	33.2%	Score 701;	DB 6;	Length 139;		
	Best Local Similarity	92.8%	Pred. No. 6.7e-50;				
	Matches 129;	Conservative	7;	Mismatches	3;	Indels	0;
						Gaps	0;
QY	48	AYKQSPNVPKPTLGASGRYEGLTARSSERKELTPNNPDIIFDEENTGADRLMTQRC	107				
Db	1	AYKQFIPNVAKTITGASGRYEGLKITRNSERKELTPNNPDIIFDEENTGADRLMTQRC	60				
QY	108	KDRUNSLAISVNNQMPGYKLTNTBGMDEDDGHHSESLTYEGRAVAVITTSDRNRKTYGLLA	167				
Db	61	KDKLNALALISVMDQMPGYKLTNTBGMDEDDGHHSESLTYEGRAVAVITTSDRNRKTYGLLA	120				
QY	168	RLAVEAGPDWVYTESKAHY	186				
Db	121	RLAVEAGPDWVYTESKAHI	139				

RESULT	9			
09M6C1				
ID	09M6C1	PRELIMINARY;	PRT;	138 AA.
AC	09M6C1;			
DT	01-NOV-1999 (TREMBLrel. 12, Created)			
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)			
DT	01-NOV-1999 (TREMBLrel. 12, Last annotation update)			
DE	SONIC HEDGEHOG PROTEIN (FRAGMENT).			
GN	SHH1			
OS	Eleutherodactylus coqui.			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;			
OC	Batrachia; Anura; Neobatrachia; Bufonidae; Leptodactylidae;			
OC	Eleutherodactylus.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	CARL T.F., RICHARDSON M.K., OLSSON L., SCHLOSSER G., KLYMKOWSKI M.W.,			
RA	HANKEN J.;			
RT	"Differences in vertebrate limb development revealed by studies of the			
RT	direct developing frog <i>E. coqui</i> ."			
RL	Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF113403; AAD23436.1; -.			
FT	NON_TER	1	1	
FT	NON_TER	138	138	
SEQ	SEQUENCE	138 AA;	15751 MW;	203E8060 CRC32;

Query Match	31.3%	Score 681	DB 13	Length 138:
Best Local Similarity	91.2%	Pred. NO. 2.8e-45		
Matches 125; Conservative	7;	Mismatches 5;	Indels 0;	Gaps 0;
QY	54	PNNVEKTLGASGREKIGARSERFKEITPPYNDIIFKDEENTGADRLMTQRCRRLNS	113	
Db	2	PNNVEKTLGASGREKIGATRSERFKEITPPYNDIIFKDEENTAADRLMTQRCRRLNS	61	
QY	114	LAISVNNQWPEVKIRVTEGWDGEGHSEESLHYGRAVDITTSRDNRKYGGLARLAVEA	173	
Db	62	LAISVNNLMFGVKIRVTEGWDGEGHSEESLHYGRAVDITTSRDNRKYGGLARLAVEA	121	
QY	174	GFDMVYTESKAHHCVS	190	
Db	122	GFDMVYTESKAHHCVS	138	
RESULT	10			
Q9YGV7				
ID	Q9YGV7	PRELIMINARY;	PRT;	150 AA.
AC	Q9YGV7;			
DT	01-MAY-1999 (TEMBLrel. 10, Created)			
DT	01-MAY-1999 (TEMBLrel. 10, Last sequence update)			
DT	01-NOV-1999 (TEMBLrel. 12, Last annotation update)			
DE	SONIC HEDGEHOG (FRAGMENT).			
OS	Ambystoma mexicanum (Xenopus laevis)			
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;			

OC Batrachia; Caudata; Salamandroidea; Ambystomatidae; Ambystoma.
RN [1]
RP SEQUENCE FROM N.A.
RA TOROK M.A., IZPILZVA-BELMONTÉ J.C., GARDINER D.M., BRYANT S.V.,
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF031480; AAD18128.1; -
HSSP: Q62226; 1VHH. 1
FT NON_TER 1 150
FT NON_TER 1 150
SQ SEQUENCE 150 AA; 16599 MW; 9356329B CRC32;

Query Match 30.6%; Score 665; DB 13; Length 150;
Best Local Similarity 82.6%; Pred. No. 6.3e-47;
Matches 123; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

QY 88 DIFKDEENTGADRLMTORCKDLNSLAISVMNOMPGVKLVTEGMDGHHSEESLHYE 147
DB 1 DIFKDEENTGADRLMTORCKDLNSLAISVMNOMPGVKLVTEGMDGHHSEESLHYE 60
QY 148 GRAVDITSDRNRKYGILARLAVEAGFDWVYSEKAVHCSSEHSAAKTGCEFPAG 207
DB 61 GRAVDITSDRNRKYGILARLAVEAGFDWVYSEKAVHCSSEHSAAKTGCEFPAG 120
QY 208 AQVPLESGARVALSAVPRGDRVLAMGEDG 236
DB 121 AKYTLHGVTAPVNDLRPRGDRVLAMGOG 149

RESULT 11
ID 096699 PRELIMINARY; PRT; 185 AA.

AC 096699;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE HEDGEHOG PROTEIN (FRAGMENT).
GN HH.
OS Junonia coenia (Peacock butterfly) (Precis coenia).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pleurogona; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Papilionoidea; Nymphalidae; Nymphalinae; Junonia.
RN [1]
RP SEQUENCE FROM N.A.
RA KEYS D.N., LEWIS D.L., SELEGUE J.E., PEARSON B.J., GOODRICH L.V.,
RA JOHNSON R.L., GATES J., SCOTT M.P., CARROLL S.B.;
RT "Recruitment of a hedgehog regulatory circuit in butterfly eyespot
evolution."
RL Science 0:0-0(1999).
DR EMBL: AF117742; AAD08931.1; -
HSSP: Q62226; 1VHH. 1
FT NON_TER 1 185
FT NON_TER 1 185
SQ SEQUENCE 185 AA; 20745 MW; 96A09B5A CRC32;

Query Match 27.4%; Score 597; DB 5; Length 185;
Best Local Similarity 65.3%; Pred. No. 2.9e-41;
Matches 111; Conservative 30; Mismatches 25; Indels 4; Gaps 3;

QY 63 ASGREGKIASSEFFKLTLPYNDITFKDEENTGADRLMTORCKDLNSLAISVMNOM 122
DB 3 ASGREGKLTDEKFDLPVNPNDIDFKDEEGADRLMTORCKDLNSLAISVMNOM 62
QY 123 PGVKLVTEGMDGHHSEESLHYEGRVADITSDRNRKYGILARLAVEAGFDWVYSE 182
DB 63 PGVKLVTEGMDGHHSEESLHYEGRVADITSDRNRKYGILARLAVEAGFDWVYSE 122
QY 183 KAHVHCYSKSHSAAKTG-GCPAGAVRLESGARVALSAVPRGDRVLA 231
DB 123 RSYHCYSKSHSAAKTG-GCPAGAVRLESGARVALSAVPRGDRVLA 169

RESULT 12

QY 09WUP6 PRELIMINARY; PRT; 129 AA.
ID 09WUP6;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE DESERT HEDGEHOG PROTEIN (FRAGMENT).
GN HH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-SPRAGUE-DAWLEY;
RA GARGES P.L., MEYER R.A., JR., BROWN C.A., PRICE D.K.;
RT "Desert hedgehog in the rat."
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF148226; AAD31927.1; -
FT NON_TER 1 129
FT NON_TER 1 129
SQ SEQUENCE 129 AA; 14578 MW; CB8B2D40 CRC32;

Query Match 23.5%; Score 510.5; DB 11; Length 129;
Best Local Similarity 72.8%; Pred. No. 1.9e-34;
Matches 91; Conservative 20; Mismatches 13; Indels 1; Gaps 1;

QY 28 CGPGR-VYSGRRRPPKLVPLAYKOFSPNVPEKTLGASGREKIASSEFFKLTLPYNYN 86
DB 5 CGPGRGVGRRRYRKQVLPILYKQFVPSMPEKTLGASGREKIASSEFFKLTLPYNYN 64
QY 87 PDIFKDEENTGADRLMTORCKDLNSLAISVMNOMPGVKLVTEGMDGHHSEESLHYE 146
DB 65 PDIFKDEENTGADRLMTORCKDLNSLAISVMNOMPGVKLVTEGMDGHHSEESLHYE 124
QY 147 EGRV 151
DB 125 EGRV 129

RESULT 13
ID 042128 PRELIMINARY; PRT; 119 AA.

AC 042128;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE SHH, PARTIAL CDS (FRAGMENT).
GN ME-SHH.
OS Oryzias latipes (Medaka fish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Atherinomorpha;
OC Cyprinodontiformes; Adrianichthyoidae; Adrianichthyidae; Oryziinae;
OC Oryzias.
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-BBRR; TISSUE-WHOLE EMBRYO;
RA ARAKI K.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB007129; BAA22368.1; -
HSSP: Q62226; 1VHH. 1
FT NON_TER 1 119
FT NON_TER 1 119
SQ SEQUENCE 119 AA; 13179 MW; 1CAE5021 CRC32;

Query Match 21.7%; Score 472; DB 13; Length 119;
Best Local Similarity 70.6%; Pred. No. 2.3e-31;
Matches 84; Conservative 17; Mismatches 18; Indels 0; Gaps 0;

QY 128 RYTEGMDGHHSEESLHYEGRVADITSDRNRKYGILARLAVEAGFDWVYSEKAVH 187
DB 128 RYTEGMDGHHSEESLHYEGRVADITSDRNRKYGILARLAVEAGFDWVYSEKAVH 187

DB 1 RTEGMDDEGHHEESLHYEGRAVDITTSRDKSKYGLTSLRLAVEAGFDWVYVESKAH1H 60
QY 188 CSVSKSHSAAKGCGCFPPAGQVRLSGARVALSAVRPGDRLVAMEGDSPTFSVLIIF 246
DB 61 CSVKAESSVAANKSGCGFPSSVTLENGTORPYKDLQPGDRLVLAADYDGNPVYTDPIMF 119
RESULT 14
QY 09YGU3 PRELIMINARY; PRT; 88 AA.
AC 09YGU3;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, last annotation update)
DE DESERT HEDGEHOG PROTEIN (FRAGMENT).
GN DH.
OS Brachydonto reatio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinidae; Rasbora; Danio.
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE; 99051425.
RA AMORES A., FORCE A., YAN Y.-L., WANG Y.-L., FRITZ A., PRINCE V.,
RA HO R., AMEMIYA C., LANGELAND J., WESTERFIELD M., EKKER M.,
RA POSTLETHWAIT J.,
RT "Zebrafish hox clusters and vertebrate genome evolution."
RL Science 282:1711-1714(1998).
DR EMBL; AF071236; AAD15931.1; -.
DR HSSP; 062226; 1VHH.
FT NON_TER 1 88
FT NON_TER 1 88
SQ SEQUENCE 88 AA; 10069 MW; 2C17DE49 CRC32;
Query Match 18.3%; Score 397; DB 13; Length 88;
Best Local Similarity 83.0%; Pred. No. 1.9e-25;
Matches 73; Conservative 5; Mismatches 10; Indels 0; Gaps 0;
QY 105 ORCKDRLNSLAISVMNQPGVKRLVTEGMDDEGHSEESLHYEGRAVDITTSRDRNKYG 164
DB 1 ORCKDCLYKLAIVMNMQPGVRLVTEAMDEGHPGSLHYEGRAVDITTSRDRTKKYG 60
QY 165 LLARLAVEAGFDWVYVESKAHVCYSK 192
DB 61 LLARLAVEAGFDWVYVESKAHVCYSKA 88
RESULT 15
QY 042441 PRELIMINARY; PRT; 80 AA.
AC 042441;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, last annotation update)
DE SONIC HEDGEHOG (FRAGMENT).
OS Oryzias latipes (Medaka fish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Atherinomorphi;
OC Cyprinodontiformes; Adrianichthyoidel; Adrianichthyidae; Oryziinae;
OC Oryzias.
RN (1)
RP SEQUENCE FROM N.A.
RX KOESTER R., STICK R., LOOSLI F., WITTBRODT J.,
RL Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL; 297019; CAB09695.1; -.
DR HSSP; 062226; 1VHH.
DR PFAW; PF01085; HH_Signal; 1.
FT NON_TER 1 80
FT NON_TER 1 80
SQ SEQUENCE 80 AA; 9180 MW; 59BE42B8 CRC32;

Query Match 17.9%; Score 389; DB 13; Length 80;
Best Local Similarity 90.0%; Pred. No. 7.3e-25;
Matches 72; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
QY 93 DEENTGADRLMTORCKDRLNSLAISVMNQPGVKRLVTEGMDDEGHSEESLHYEGRAVD 152
DB 1 DEENTGADRLMTORCKDRLNSLAISVMNQPGVKRLVTEGMDDEGHHEEFELHYEGRAVD 60
QY 153 ITTSRDRNKYGLRLAVE 172
DB 61 ITTSRDRSKYGLTSLRLAVE 80

Search completed: June 5, 2000, 08:19:30
Job time: 184 sec

Thu Jun 8 15:52:41 2000

us-08-900-220-16.rsp

Page 7

THIS PAGE BLANK (USPTO)

DB 121 QMPGKLTGTEWDEDESESLHTEGRAVDITTSDRNRKNGYLARLAVAGDWMY 180
QY 181 ESKAHVHGSVSESHSAATGCGCPAGAOVRLSEGARVALSAVRPGDRLVLMGEGSPTE 240
DB 181 ESKAHVHGSVSESHSAATGCGCPAGAOVRLSEGARVALSAVRPGDRLVLMGEGSPTE 240
QY 241 SVLLFLDREPRRLRAFOVIEETODPPRLALTPAHLFTADNHTPEPARFRATFASHVP 300
DB 241 SVLLFLDREPRRLRAFOVIEETODPPRLALTPAHLFTADNHTPEPARFRATFASHVP 300
QY 301 GOYVYVAGVPGLOPARVAAVSTHVALGAYAPLTKHGTIVVEDVVASCPAAVADHHLAOLA 360
DB 301 GOYVYVAGVPGLOPARVAAVSTHVALGAYAPLTKHGTIVVEDVVASCPAAVADHHLAOLA 360
QY 361 FMPRLRFLSHLAWGSMTPGEGVHWYPOLLYRLGRLLLEEGSFHPLGMSGAGS 411
DB 361 FMPRLRFLSHLAWGSMTPGEGVHWYPOLLYRLGRLLLEEGSFHPLGMSGAGS 411
RESULT 2
W97763 standard; Protein: 411 AA.
ID W97763
AC W97763
DE 21-MAY-1999 (first entry)
DE Human Indian hedgehog (Ihh) protein.
KW Indian hedgehog; Ihn protein; human; dopaminergic; GABA-nergic;
KW ptc therapeutic; patched; signal transduction; Parkinson's disease;
KW Huntington's disease; amyotrophic lateral sclerosis;
KW cerebral ischemia; hypoxia; neuroprotective; therapy.
OS Homo sapiens.
PN W09804775-A2.
PD 04-FEB-1999.
PF 24-JUL-1998; U15419.
PR 24-JUL-1997; US-900220.
PA (ONTO-) ONTOGENY INC.
PI Mahanthappa NK, Miao N, Pang K, Wang M;
WPI: 99-142578/12.
DR N-PSDB: X07269.
PT Increasing the survival of neuronal, dopaminergic and GABA-nergic
PT cells - by using a ptc therapeutic such as a protein kinase
PT inhibitor, or an agent derived from hedgehog polypeptides, useful in
PT the treatment of Parkinson's disease.
PS Claim 30; Page 95-97; 138pp; English.
CC This polypeptide is human Indian hedgehog protein (Ihh). The
CC invention is based on the finding that hedgehog proteins are useful
CC as protective agents in the treatment and prophylaxis of
CC neurodegenerative disorders resulting from the loss of dopaminergic
CC and/or GABA-nergic neurons, or the general loss of tissue from the
CC substantia nigra. Exemplary disorders include Parkinson's disease,
CC Huntington's disease (both claimed); amyotrophic lateral sclerosis
CC and cerebral ischemia. The invention relates to hedgehog
CC therapeutics (i.e. hedgehog polypeptides and gene therapy constructs
CC e.g. constructs encoding recombinant hedgehog polypeptides and
CC trans-activation constructs for altering hedgehog gene regulatory
CC sequences) and ptc therapeutics (i.e. agents which mimic the effect
CC of naturally occurring hedgehog proteins on patched signaling)
CC that are effective in both human and animal subjects. A bioactive
CC polypeptide comprising amino acids 28-202 of human Ihn is preferred.
CC The products can also be used for the maintenance of differentiated
CC neurons in cultures, and to enhance the implantation of such
CC neuronal cells in an animal. They can be used to prevent or treat
CC neurodegenerative conditions arising from the use of certain drugs,
CC and in the prevention and/or treatment of hypoxia, e.g. as a
CC neuroprotective agent.
CC Sequence 411 AA:
QY 1 MSPARLRPLRHLFCVLLLVLPAAWGCGPGRVSGRRPRPKVLPAIKQSPNPPEKT 60
Query Match 100.0%; Score 2175; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 1,1e-213;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 MSPARLRPLRHLFCVLLLVLPAAWGCGPGRVSGRRPRPKVLPAIKQSPNPPEKT 60
QY 61 LGASRYEGKTLARSSERKELTPNPNPDIIFKDEBNTGADRLMORCKDRNLNSLAISVN 120
DB 61 LGASRYEGKTLARSSERKELTPNPNPDIIFKDEBNTGADRLMORCKDRNLNSLAISVN 120
QY 121 QMPGKLTGTEWDEDESESLHTEGRAVDITTSDRNRKNGYLARLAVAGDWMY 180
DB 121 QMPGKLTGTEWDEDESESLHTEGRAVDITTSDRNRKNGYLARLAVAGDWMY 180
QY 181 ESKAHVHGSVSESHSAATGCGCPAGAOVRLSEGARVALSAVRPGDRLVLMGEGSPTE 240
DB 181 ESKAHVHGSVSESHSAATGCGCPAGAOVRLSEGARVALSAVRPGDRLVLMGEGSPTE 240
QY 241 SVLLFLDREPRRLRAFOVIEETODPPRLALTPAHLFTADNHTPEPARFRATFASHVP 300
DB 241 SVLLFLDREPRRLRAFOVIEETODPPRLALTPAHLFTADNHTPEPARFRATFASHVP 300
QY 301 GOYVYVAGVPGLOPARVAAVSTHVALGAYAPLTKHGTIVVEDVVASCPAAVADHHLAOLA 360
DB 301 GOYVYVAGVPGLOPARVAAVSTHVALGAYAPLTKHGTIVVEDVVASCPAAVADHHLAOLA 360
QY 361 FMPRLRFLSHLAWGSMTPGEGVHWYPOLLYRLGRLLLEEGSFHPLGMSGAGS 411
DB 361 FMPRLRFLSHLAWGSMTPGEGVHWYPOLLYRLGRLLLEEGSFHPLGMSGAGS 411
RESULT 3
Y05516 standard; Protein: 411 AA.
ID Y05516
AC Y05516
DE 05-JUL-1999 (first entry)
DE Human Indian hedgehog protein Ihn.
KW Indian hedgehog; Ihn protein; human; hedgehog therapeutic;
KW ptc therapeutic; patched; signal transduction; muscle atrophy;
KW cachexia; muscular myopathy; myoblastic sarcoma; therapy.
OS Homo sapiens.
PN W09910004-A2.
PD 04-MAR-1999.
PF 28-AUG-1998; U17922.
PR 29-AUG-1997; US-057394.
PA (ONTO-) ONTOGENY INC.
PI Bladgen CS, Currie PD, Hughes SM, Ingham PW;
WPI: 99-243557/20.
DR N-PSDB: X25104.
PT A new method to regulate muscle growth
PT disclosure: Page 120-121; 130pp; English.
PS The present sequence is human Indian hedgehog protein Ihn. The
PS invention relates to a method for modulating the formation and/or
PS maintenance of muscle tissue by ectopically contacting muscle
PS cells, especially muscle stem/progenitor cells, in vitro or in
PS vivo, with a hedgehog therapeutic (i.e. hedgehog polypeptides and
PS gene therapy constructs) or ptc therapeutic (i.e. a small organic
PS molecule that mimics the effect of hedgehog proteins on patched
PS signaling, or activates or potentiates patched signaling) in an
PS amount effective to alter the growth state of the treated cells.
PS Also claimed is a method for treatment or prevention of disorders
PS of, or surgical or cosmetic repair of, such muscle tissues, by
PS administering a hedgehog polypeptide or ptc therapeutic. The
PS disorder may be muscle atrophy, in particular skeletal muscle
PS atrophy or cardiac muscle atrophy, cachexia, or muscular myopathy
PS (all claimed). The hedgehog polypeptide or ptc therapeutic can
PS inhibit growth of myoblastic-derived tissue or provide treatment of
PS hyperplastic or neoplastic growth of muscle tissue such as in
PS myoblastic sarcoma (also claimed). The hedgehog therapeutic
PS preferably comprises at least a bioactive extracellular portion of
PS a hedgehog protein (see Y05510-19) encoded by a vertebrate hedgehog
PS gene (see X25098-107), especially a human hedgehog gene.
PS Sequence 411 AA:
QY 1 MSPARLRPLRHLFCVLLLVLPAAWGCGPGRVSGRRPRPKVLPAIKQSPNPPEKT 60
Query Match 100.0%; Score 2175; DB 1; Length 411;

Best Local Similarity 100.0%; Pred. NO. 1.1e-213;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MSPARLPRRLFFLCYUULLLLLVPAAMCCGGGRVYSSRRPRKLYPLATQFSPNNPEKT	60
Db	1	MSPARLPRRLFFLCYUULLLLLVPAAMCCGGGRVYSSRRPRKLYPLATQFSPNNPEKT	60
QY	61	LGASGRYEGKTIARSSSEFFKELTPNYNDIIFKDEENTGADRLMTQRCXDRINLSAISVM	120
Db	61	LGASGRYEGKTIARSSSEFFKELTPNYNDIIFKDEENTGADRLMTQRCXDRINLSAISVM	120
QY	121	QMPGVKLRYTEGMDGDHNSFESLHYSGRAVDITTSRDNRKYGILLRLAYAGFDMVUY	180
Db	121	QMPGVKLRYTEGMDGDHNSFESLHYSGRAVDITTSRDNRKYGILLRLAYAGFDMVUY	180
QY	181	ESKANHCSYKSEHSAAKTCGCPACAGOVYLESGAVALSAPVGRGVYUAMGDSGPTF	240
Db	181	ESKANHCSYKSEHSAAKTCGCPACAGOVYLESGAVALSAPVGRGVYUAMGDSGPTF	240
QY	241	SDVLIFFLDREPHRYARAFQVETQDPRRLALTAPRIHLFTADNHTEPARFRATASVQF	300
Db	241	SDVLIFFLDREPHRYARAFQVETQDPRRLALTAPRIHLFTADNHTEPARFRATASVQF	300
QY	301	GOYUUVAGVPELOPARVAANSTHVALGATAYARLIKNGSLVVEDVVASCFPAUADNHLAOLA	360
Db	301	GOYUUVAGVPELOPARVAANSTHVALGATAYARLIKNGSLVVEDVVASCFPAUADNHLAOLA	360
QY	361	FWPRLRHLNSLAWGSMTPGEGVHNMPOLLRYLGRILLTEGSHRPLGMSGAGS	411
Db	361	FWPRLRHLNSLAWGSMTPGEGVHNMPOLLRYLGRILLTEGSHRPLGMSGAGS	411

RESULT	4
W94470	
ID	W94470 standard; Protein; 411 AA

29-APR-1999 (first entry)
Mouse Ihh hedgehog protein sequence.
KW Patched; hedgehog; plc therapeutic; neuroprotective; neuronal cell;
KW brain infarction; cerebral infarction; transient ischaemic attack;
KW stroke; cerebral infarct volume; spinal cord; oedema; trauma;
KW haemorrhage; encephalomyelitis; coronary bypass; cerebral hypoxia.
OS Mus sp.
PN M09900117-A2.
PD 07-JAN-1999.
PF 26-JUN-1998; U13387.
PR 27-JUN-1997; US-883656.
PA (ONTO-) ONTOGENY INC 
PI Mahanthappa, NK.
DR WPI; 99-095458/08.
N-PSDS: X16184.
PT Method for limiting damage to neurons caused by ischaemic or epoxic
PT conditions - is used for the treatment and prevention of e.g.
PT cerebral infarction, stroke and transient ischaemic attacks
PT Disclosure; Page 66-68; 104pp; English.
PS A method has been developed for limiting the damage to neuronal cells by
PS ischaemic or epoxic conditions by administering a plc (patched) Damage
CC therapeutic agent to reduce cerebral infarct volume (CIV). Damage to
CC neuronal cells can also be limited by administering a gene activation
CC construct which recombines with the genomic hedgehog gene to provide a
CC heterologous transcription regulator linked to the coding region of this
CC gene. Administration of the plc therapeutic agent is used to protect
CC cerebral tissues against ischaemic injury; to treat cerebral infarct or
CC ischaemia, stroke (thrombotic or embolic) and transient ischaemic
CC attacks. It may also be used as a prophylactic in many other cases of
CC injury to the brain or spinal cord, oedema caused by trauma, haemorrhages
CC and encephalomyelitis, or in conjunction with (coronary bypass) surgery.
CC Treatment (which may be prophylactic) is used where ischaemic/epoxic
CC conditions may cause cerebral hypoxia, or progressive loss of neurons
CC due to oxygen depletion, including in patients with hypotension. The
CC treatment reduces CIV by at least 25, particularly at least 70,%. The
CC present sequence represents a hedgehog sequence given in the present
CC invention.

Sequence 411 AA;

Query Match	93.88;	Score 2041;	DB 1;	Length 411;
Best Local Similarity	94.68;	Pred. No. 5.3e-200;		
Matches	389;	Conservative	10;	Mismatches 12;
				Indels 0;
				Gaps 0;

QY	1	MSPALRRLTFFCJVLILLVLPAAACGGRVVSGRRPRKLYPLATQFSPNPEKT	60
Db	1	MSPALRRLRFLFCJLILLVLPAAACGGRVVSGRRPRKLYPLATQFSPNPEKT	60
QY	61	LGASRGYEGKTIARSSEREKELTPNYNDIIFKDEENTGADRLMTQRCXDRILNSLAISVM	120
Db	61	LGASRGYEGKTIARSSEREKELTPNYNDIIFKDEENTGADRLMTQRCXDRILNSLAISVM	120
QY	121	QMPGYKLAVTGMDDEGHHSESLHYEGRAVDITTSRDNRKIGLLRLAVEAGFDWYU	180
Db	121	QMPGYKLAVTGMDDEGHHSESLHYEGRAVDITTSRDNRKIGLLRLAVEAGFDWYU	180
QY	181	ESKAHVHGSVESEHSAAKTGCCPCAGQVRLSEGAVALSAVPPGDRVLAMGDSGPTF	240
Db	181	ESKAHVHGSVESEHSAAKTGCCPCAGQVRLSEGAVALSAVPPGDRVLAMGDSGPTF	240
QY	241	SDVLFIDREERHRLRAFOVITEOPPRRLATLPAHLTLFTADNHTEPARERATFASHVOP	300
Db	241	SDVLFIDREERHRLRAFOVITEOPPRRLATLPAHLTLFTADNHTEPARERATFASHVOP	300
QY	301	GOYUUVAGVPGLOPARVAAVSTHYALGATAPLTNGHGLVVEDVYASCPFAAVADHNLQOLA	360
Db	301	GOYUUVAGVPGLOPARVAAVSTHYALGATAPLTNGHGLVVEDVYASCPFAAVADHNLQOLA	360
QY	361	FMPRLHFLSLAMGSMWTPEGVHWYRPLLRYLGRLLTREGSGHPLRGMGGAAS	411
Db	361	FMPRLHFLSLAMGSMWTPEGVHWYRPLLRYLGRLLTREGSGHPLRGMGGAAS	411

RESULT	5
W97767	
ID	W97767 standard; Protein; 411 AA.

DE 21-MAY-1999 (first entry)
 DE Mouse Indian hedgehog (Ihh) protein.
 KM Indian hedgehog; Ihh protein; mouse; dopaminergic; GABA-nergic;
 KM ptc therapeutic; patched; signal transduction; Parkinson's disease;
 KM Huntington's disease; amyotrophic lateral sclerosis;
 KM cerebral ischemia; hypoxia; neuroprotective; therapy.
 OS Mus sp.
 PN W09904775-A2.
 PD 04-FEB-1999.
 PF 24-JUL-1998; U15419.
 PR 24-JUL-1997; US-900220.
 PA (ONTO-) ONTOGENY INC.
 PI Mahanthappa NK, Miao N, Pang K, Wang M;
 PI WPI: 99-142578/12.
 DR N-PSDB: X07273.
 PT Increasing the survival of neuronal, dopaminergic and GABA-nergic
 PT cells by using a ptc therapeutic such as a protein kinase
 PT inhibitor, or an agent derived from hedgehog polypeptides, useful in
 PT the treatment of Parkinson's disease.
 PS Disclosure; Page 87-88; 13pp; English.
 CC This polypeptide is mouse Ihh Indian hedgehog protein. The
 CC invention is based on the finding that hedgehog proteins are useful
 CC as protective agents in the treatment and prophylaxis of
 CC neurodegenerative disorders resulting from the loss of dopaminergic
 CC and/or GABA-nergic neurons, or the general loss of tissue from the
 CC substantia nigra. Exemplary disorders include Parkinson's disease,
 CC Huntington's disease (both claimed), amyotrophic lateral sclerosis
 CC and cerebral ischemia. The invention relates to hedgehog
 CC therapeutics (i.e. hedgehog polypeptides and gene therapy constructs
 CC e.g. constructs encoding recombinant hedgehog polypeptides and
 CC trans-activation constructs for altering hedgehog gene regulatory
 CC sequences) and ptc therapeutics (i.e. agents which mimic the effect
 CC of naturally occurring hedgehog proteins on patched signalling)

CC that are effective in both human and animal subjects. Human Ihh
 CC and Dhh polypeptides (see W97/63-64) are preferred. The products
 CC can also be used for the maintenance of differentiated neurons in
 CC cultures, and to enhance the implantation of such neuronal cells in
 CC an animal. They can be used to prevent or treat neurodegenerative
 CC conditions arising from the use of certain drugs, and in the
 CC prevention and/or treatment of hypoxia, e.g. as a neuroprotective
 CC agent.
 SQ Sequence 411 AA;

Query Match 93.8%; Score 2041; DB 1; Length 411;
 Best Local Similarity 94.6%; Pred. No. 35.3e-200;
 Matches 389; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 1 MSPARLRRLHFCVLVLLLVPAAMGCGPGRVGSRRRPRKLYPLAYKQFSPNVEKT 60
 DB 1 MSPARLRRLHFCVLVLLLVPAAMGCGPGRVGSRRRPRKLYPLAYKQFSPNVEKT 60
 QY 61 LGASGRYEGKIARSSERFELTPNPNPDIIFKDEENTGADRLMTORCKDRLSLAISVMN 120
 DB 61 LGASGRYEGKIARSSERFELTPNPNPDIIFKDEENTGADRLMTORCKDRLSLAISVMN 120
 QY 121 QMPVKRLVTEGMDGDHSESLHYEGRAVDITSDNRNRYGLARLAVAGFDWYY 180
 DB 121 QMPVKRLVTEGMDGDHSESLHYEGRAVDITSDNRNRYGLARLAVAGFDWYY 180
 QY 181 ESKAHVHCYSVSEHSAAKTGCGCPAGAOVRLSEGAVALSAVRPGDRILAMGEGSPTE 240
 DB 181 ESKAHVHCYSVSEHSAAKTGCGCPAGAOVRLSEGAVALSAVRPGDRILAMGEGSPTE 240
 QY 241 SDVLIIFLDREPHRLARFOVLETODPPRLALTPAHLFTADNHTPEARAFATASHVOP 300
 DB 241 SDVLIIFLDREPHRLARFOVLETODPPRLALTPAHLFTADNHTPEARAFATASHVOP 300
 QY 301 GOYLVAGVGPGLPARAAVSTHVALGSAVPLTRHGTLVEDVVAASCPAAVADHHLAQLA 360
 DB 301 GOYLVAGVGPGLPARAAVSTHVALGSAVPLTRHGTLVEDVVAASCPAAVADHHLAQLA 360
 QY 361 FWPRLRFLHSLAMGSWTPGEGVHWYPOLLYRLGRLLLEGSFHPPLGMSGAGS 411
 DB 361 FWPRLRFLHSLAMGSWTPGEGVHWYPOLLYRLGRLLLEGSFHPPLGMSGAGS 411

RESULT 6
 Y05512
 ID Y05512 standard; Protein; 411 AA.
 AC Y05512;
 DT 05-JUL-1999 (first entry)
 DE Mouse Indian hedgehog protein Ihh.
 KW Indian hedgehog; Ihh protein; mouse; hedgehog therapeutic;
 KM ptc therapeutic; patched; signal transduction; muscle atrophy;
 OS cecchiella; muscular myopathy; myoblastic sarcoma; therapy.
 PN Mus sp.
 MO9910004-A2.
 PD 04-MAR-1999.
 PF 28-AUG-1998; U17922.
 PR 29-AUG-1997; US-057394.
 PA (ONTO-) ONTOGENY INC.
 PI Bladen CS, Currie PD, Hughes SM, Ingham PW;
 DR WPI: 99-243557/20.
 DR N-PSDB; X25100.
 PT A new method to regulate muscle growth
 PS Disclosure; Page 114-115; 130pp; English.
 CC The present sequence is mouse Indian hedgehog protein Ihh. The
 CC invention relates to a method for modulating the formation and/or
 CC maintenance of muscle tissue by ectopically contacting muscle
 CC cells, especially muscle stem/progenitor cells, in vitro or in
 CC vivo, with a hedgehog therapeutic (i.e. hedgehog polypeptides and
 CC gene therapy constructs) or ptc therapeutic (i.e. a small organic
 CC molecule that mimics the effect of hedgehog proteins on patched
 CC signaling, or activates or potentiates patched signaling) in an
 CC amount effective to alter the growth state of the treated cells.

CC Also claimed is a method for treatment or prevention of disorders
 CC of, or surgical or cosmetic repair of, such muscle tissues, by
 CC administering a hedgehog polypeptide or ptc therapeutic. The
 CC disorder may be muscle atrophy, in particular skeletal muscle
 CC atrophy or cardiac muscle atrophy, cachexia, or muscular myopathy
 CC (all claimed). The hedgehog polypeptide or ptc therapeutic can
 CC inhibit growth of myoblastic-derived tissue to provide treatment of
 CC hyperplastic or neoplastic growth of muscle tissue such as in
 CC myoblastic sarcoma (also claimed). The hedgehog therapeutic
 CC preferably comprises at least a bioactive extracellular portion of
 CC a hedgehog protein (see Y05510-19) encoded by a vertebrate hedgehog
 CC gene (see X25098-107), especially a human hedgehog gene.
 SQ Sequence 411 AA;

Query Match 93.8%; Score 2041; DB 1; Length 411;
 Best Local Similarity 94.6%; Pred. No. 5.3e-200;
 Matches 389; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 1 MSPARLRRLHFCVLVLLLVPAAMGCGPGRVGSRRRPRKLYPLAYKQFSPNVEKT 60
 DB 1 MSPARLRRLHFCVLVLLLVPAAMGCGPGRVGSRRRPRKLYPLAYKQFSPNVEKT 60
 QY 61 LGASGRYEGKIARSSERFELTPNPNPDIIFKDEENTGADRLMTORCKDRLSLAISVMN 120
 DB 61 LGASGRYEGKIARSSERFELTPNPNPDIIFKDEENTGADRLMTORCKDRLSLAISVMN 120
 QY 121 QMPVKRLVTEGMDGDHSESLHYEGRAVDITSDNRNRYGLARLAVAGFDWYY 180
 DB 121 QMPVKRLVTEGMDGDHSESLHYEGRAVDITSDNRNRYGLARLAVAGFDWYY 180
 QY 181 ESKAHVHCYSVSEHSAAKTGCGCPAGAOVRLSEGAVALSAVRPGDRILAMGEGSPTE 240
 DB 181 ESKAHVHCYSVSEHSAAKTGCGCPAGAOVRLSEGAVALSAVRPGDRILAMGEGSPTE 240
 QY 241 SDVLIIFLDREPHRLARFOVLETODPPRLALTPAHLFTADNHTPEARAFATASHVOP 300
 DB 241 SDVLIIFLDREPHRLARFOVLETODPPRLALTPAHLFTADNHTPEARAFATASHVOP 300
 QY 301 GOYLVAGVGPGLPARAAVSTHVALGSAVPLTRHGTLVEDVVAASCPAAVADHHLAQLA 360
 DB 301 GOYLVAGVGPGLPARAAVSTHVALGSAVPLTRHGTLVEDVVAASCPAAVADHHLAQLA 360
 QY 361 FWPRLRFLHSLAMGSWTPGEGVHWYPOLLYRLGRLLLEGSFHPPLGMSGAGS 411
 DB 361 FWPRLRFLHSLAMGSWTPGEGVHWYPOLLYRLGRLLLEGSFHPPLGMSGAGS 411

RESULT 7
 R77343
 ID R77343 standard; Protein; 336 AA.
 AC R77343;
 DT 18-MAR-1996 (first entry)
 DE Mouse Indian hedgehog protein.
 KW Mouse; Indian hedgehog protein; probe; primer; diagnostic;
 KM nervous system disorder; gene therapy; antibody.
 OS Mus musculus.
 PN MO9518856-A1.
 PD 13-UTL-1995.
 PF 30-DEC-1994; U14992.
 PR 30-DEC-1993; US-176427.
 PR 14-DEC-1994; US-356060.
 PA (HARD) HARVARD COLLEGE.
 PI (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
 DR Ingham PW, McMahon AP, Tabin CJ;
 DR WPI: 95-255060/33.
 DR N-PSDB; O91640.
 PT Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful
 PT to treat degenerative nervous system disorder(s) and in gene
 PT therapy.
 PS Claim 17; Page 137-38; 210pp; English.
 CC The sequence represents a mouse Indian hedgehog protein, homologous
 CC to a Drosophila hedgehog protein (R77337), and is encoded by a cDNA

isolated by low stringency screening of a mouse genome DNA library and screening of an 8.5 day post coitum cDNA library. Probes and primers derived from hedgehog sequences may be used as diagnostic agents for neuromuscular, autonomic or central nervous system disorders, and the gene may also be used in gene therapy. Antibodies generated from the protein may be used as therapeutic or research reagents.

Sequence 336 AA;

Query Match 77.5%; Score 1686; DB 1; Length 336;
Best Local Similarity 94.9%; Pred. No. 6,6e-164;
Matches 319; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY 76 ERFKELPNYNDILFKDEENTGADRLMTQRCRDLNSLAISVMQMPGVKLRTGMD 135
D 1 ERFKELPNYNDILFKDEENTGADRLMTQRCRDLNSLAISVMQMPGVKLRTGMD 60
QY 136 DGHSESLHTEGRAVDITSDRDNKYGGLARLAVENGFWMYTESKAHVCSKSEHS 195
D 61 DGHSESLHTEGRAVDITSDRDNKYGGLARLAVENGFWMYTESKAHVCSKSEHS 120
QY 196 AAKTGGCFPAGQVRLSESGARVALSAVRPGDVLAMGEDSPFSDVLIFLDRPHRLR 255
D 121 AAKTGGCFPAGQVRLSESGARVALSAVRPGDVLAMGEDSPFSDVLIFLDRPHRLR 180
QY 256 AFQVETODPPRLALTPAHLIFTDNHTEPARFATFASHVQPGQYVLAVAGVGLQPA 315
D 181 AFQVETODPPRLALTPAHLIFTDNHTEPARFATFASHVQPGQYVLAVAGVGLQPA 240
QY 316 RVAANSTHVALGAVAPLRFHGLTVEDVVASCFEAVADHHLAQLAFWPLRFLHSLAMGSM 375
D 241 RVAANSTHVALGAVAPLRFHGLTVEDVVASCFEAVADHHLAQLAFWPLRFLHSLAMGSM 300
QY 376 TPGEVHWYPOLLYRLGRLLLESGSFHPLGMSGAGS 411
D 301 TPGEVHWYPOLLYRLGRLLLESGSFHPLGMSGAGS 336

RESULT 8
R77344
ID R77344 standard; protein; 312 AA.
AC R77344:
DE 18-MAR-1996 (first entry)
DT Human Indian hedgehog protein.
KW Human; Indian hedgehog protein; probe; primer; diagnostic;
KW nervous system disorder; gene therapy; antibody.
OS Homo sapiens.
PN M09518856-A1.
PD 13-JUL-1995.
PF 30-DEC-1994; U14992.
PR 30-DEC-1993; US-176427.
PR 14-DEC-1994; US-356060.
PA (HARD) HARVARD COLLEGE.
PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
PI Ingham PW, Mcmanus AP, Tabin CJ;
DR WPI: 95-255060/33.
DR N-PSDB: 091641.
PT Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful
PT to treat degenerative nervous system disorder(s) and in gene
PT therapy.
PS Claim 17; Page 146-47; 210pp; English.
CC The sequence represents a mouse Indian hedgehog protein, homologous
CC to a Drosophila hedgehog protein (R77337), and is encoded by a cDNA
CC isolated by screening of a human fetal lung 5'-stretch plus cDNA
CC library. Probes and primers derived from hedgehog sequences may be
CC used as diagnostic agents for neuromuscular, autonomic or central
CC nervous system disorders, and the gene may also be used in gene
CC therapy. Antibodies generated from the protein may be used as
CC therapeutic or research reagents.
SQ Sequence 312 AA;

Query Match 75.7%; Score 1646; DB 1; Length 312;
Best Local Similarity 100.0%; Pred. No. 7.1e-160;
Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 RLMTQRCRDLNSLAISVMQMPGVKLRTGMDGHHSESLHTEGRAVDITSDRDR 160
D 101 RLMTQRCRDLNSLAISVMQMPGVKLRTGMDGHHSESLHTEGRAVDITSDRDR 61
QY 161 NKYGGLARLAVENGFWMYTESKAHVCSKSEHSAAKTGGCFPAGQVRLSESGARVAL 220
D 62 NKYGGLARLAVENGFWMYTESKAHVCSKSEHSAAKTGGCFPAGQVRLSESGARVAL 121
QY 221 SAVRPGDVLAMGEDSPFSDVLIFLDRPHRLRATVITODPPRLALTPAHLIFTA 280
D 122 SAVRPGDVLAMGEDSPFSDVLIFLDRPHRLRATVITODPPRLALTPAHLIFTA 181
QY 281 DNHTEPARFATFASHVQPGQYVLAVAGVGLQPARVAANSTHVALGAVAPLRFHGLTV 340
D 182 DNHTEPARFATFASHVQPGQYVLAVAGVGLQPARVAANSTHVALGAVAPLRFHGLTV 241
QY 341 EDVVASCFEAVADHHLAQLAFWPLRFLHSLAMGSMTPGEVHWYPOLLYRLGRLLLESGS 400
D 242 EDVVASCFEAVADHHLAQLAFWPLRFLHSLAMGSMTPGEVHWYPOLLYRLGRLLLESGS 301
QY 401 FHPGLMSGAGS 411
D 302 FHPGLMSGAGS 312

RESULT 9
W61488
ID W61488 standard; protein; 437 AA.
AC W61488:
DE 20-OCT-1998 (first entry)
DT Mouse sonic hedgehog (shh) protein.
KW Hedgehog polypeptide; neuronal cell proliferation; zebrafish; shh;
KW cholesterol biosynthesis; pituitary gland gene expression; mouse;
KW liggy-winkle hedgehog protein; twhh; sonic hedgehog; major limb trauma;
KW neuronal degeneration; nerve-sparing agent; Smith-Lemli-Opitz syndrome.
OS Mus sp.
PN M09830576-A1.
PD 16-JUL-1998.
PF 07-OCT-1997; U15753.
PR 02-OCT-1997; US-061823.
PR 07-OCT-1996; US-729743.
PA (UYUO) UNIV JOHNS HOPKINS SCHOOL OF MEDICINE.
PI Beachy PA, Porter JA;
DR WPI: 98-399053/34.
PT New hedgehog-derived polypeptide(s) - used to develop products for
PT modulating proliferation or differentiation of neuronal cells,
PT cholesterol biosynthesis or transport or expression of pituitary
PT gland gene(s)
PS Claim 13; Pages 153-154; 210pp; English.
CC This represents a mouse sonic hedgehog (shh) protein sequence. The
CC invention provides methods and compounds for modulating proliferation or
CC differentiation of neuronal cells, cholesterol biosynthesis or transport
CC or expression of pituitary gland genes. The method for affecting
CC cholesterol biosynthesis or transport in a cell comprises contacting a
CC cell with a compound that affects hedgehog, thereby affecting cholesterol
CC biosynthesis or transport. The methods for inhibiting the neural inducing
CC activity of a hedgehog polypeptide in cells, and for inducing pituitary
CC gland gene expression utilizes sequences selected from a zebrafish
CC liggy-winkle hedgehog (twhh) protein (W61485), a zebrafish sonic hedgehog
CC (shh) protein (W61486), a chicken shh protein (W61487) or a mouse shh
CC protein (W61488). The products and methods provide for compounds which
CC can affect hedgehog activity. They can be used for treating disorders
CC which arise from neuronal degeneration or abnormal function. They can
CC also be used as nerve-sparing agents or in restoring or promoting
CC appropriate patterning during the healing of major limb trauma. They can
CC also be used for treating Smith-Lemli-Opitz syndrome. The products can
CC also be used for detection and diagnosis.
SQ Sequence 437 AA;

PN WO910004-A2.
 PD 04-MAR-1999.
 PR 28-AUG-1998; U17922.
 PR 29-AUG-1997; US-057394.
 PI (ONTO-) ONTOGENY INC.
 PI Bladgen CS, Currie PD, Hughes SM, Ingham PW;
 DR WPI: 99-243557/20.
 DR N-PSDB: X25101.
 PT A new method to regulate muscle growth.
 PS Disclosure: Page 115-116; 130pp; English.
 CC The present sequence is mouse sonic hedgehog protein Shh. The
 CC invention relates to a method for modulating the formation and/or
 CC maintenance of muscle tissue by ectopically contacting muscle
 CC cells, especially muscle stem/progenitor cells, in vitro or in
 CC vivo, with a hedgehog therapeutic (i.e. hedgehog polypeptides and
 CC gene therapy constructs) or ptc therapeutic (i.e. a small organic
 CC molecule that mimics the effect of hedgehog proteins on patched
 CC signaling, or activates or potentiates patched signaling) in an
 CC amount effective to alter the growth state of the treated cells.
 CC Also claimed is a method for treatment or prevention of disorders
 CC of, or surgical or cosmetic repair of, such muscle tissues, by
 CC administering a hedgehog polypeptide or ptc therapeutic. The
 CC disorder may be muscle atrophy, in particular skeletal muscle
 CC atrophy or cardiac muscle atrophy, cachexia, or muscular myopathy
 CC (all claimed). The hedgehog polypeptide or ptc therapeutic can
 CC inhibit growth of myoblastic-derived tissue to provide treatment of
 CC hyperplastic or neoplastic growth of muscle tissue such as in
 CC myoblastic sarcoma (also claimed). The hedgehog therapeutic
 CC preferably comprises at least a bioactive extracellular portion of
 CC a hedgehog protein (see Y0510-19) encoded by a vertebrate hedgehog
 CC gene (see X25098-107), especially a human hedgehog gene.
 SO Sequence 437 AA:

Query Match 58.7%; Score 1276.5; DB 1; Length 437;
 Best Local Similarity 60.8%; Pred. No. 6e-122;
 Matches 261; Conservative 40; Mismatches 91; Indels 37; Gaps 8;

13 CLVLL-LLVPAAMGCGPRGVGRRPRKLVPLAYKQFSPVNPVKTKGASGRYG 69
 DB 8 CLVLIASSLLVCPGL-ACGPGRGFC-KRRHPRKLTPLAYKQFIPVNAKTIAGSGRYEG 65
 QY 70 KARSSEKREKELTPNYNPDIIFKDEENTGADRLMTORCKDRKLSLAISVMNOMPQKLYR 129
 DB 66 KTRNSERKELTPNYNPDIIFKDEENTGADRLMTORCKDRKLSLAISVMNOMPQKLYR 125
 QY 130 TEGWDEGHSESLHYEGRAVDITTSDDRNKYGILALAYEAGFDWYYSKAVHCS 189
 DB 126 TEGWDEGHSESLHYEGRAVDITTSDDRNKYGILALAYEAGFDWYYSKAVHCS 185
 QY 190 VASEHSAAKTGGCPAGAOVRLESGARVALSAVRPGDVLAMGEDGSTFSDVLIFFDR 249
 DB 186 VAENSVAAKSGGCPGSAVTLHEGGCTLVLDLRGDDVLAADOGRLYDFFLFFDR 245
 QY 250 EPHRAAFQVITOPPRRLATLPAHLLETA-DNHTEPARFRATFASHVQGVVLVAG 308
 DB 246 DSGAKKVFVITLTPREKLLTLAAHLFVAPRNDGSPFGSALFASVVRGQRYVYA 305
 QY 309 VEG---LQPARVAANS-THVALGAYAPLTGKGTLVEDVVASCPAAVADHHLAQLAFWP 363
 DB 306 ERGGDRRLPLPAVHSTLREEAGAVAPLTAGHTLILNNVLSAIVIEHSMARAFAP 365
 QY 364 LRLFLSL-----AMGSMTPGEGVHHYPOLLYLGRLLLEE 398
 DB 366 FLTAHLAALAPARTDGGGSGIPAASQATEARGA-EPTAGIHYSQLLHYHTWMLDS 424
 QY 399 GSFHPLGMS 407
 DB 425 ETMHPLGMA 433

RESULT 14
 W61486

ID W61486 standard; Protein; 418 AA.
 AC W61486;
 DE 20-OCT-1998 (first entry)
 DT Zebrafish sonic hedgehog (shh) protein.
 KW Hedgehog polypeptide; neuronal cell proliferation; zebrafish;
 KW cholesterol biosynthesis; pituitary gland gene expression; shh;
 KW tiggly-winkle hedgehog protein; twbn, sonic hedgehog; major limb trauma;
 KW neuronal degeneration; nerve-sparing agent; Smith-Lemli-Opitz syndrome.
 OS Brachydanio rerio.
 PN WO9830576-A1.
 PD 16-JUL-1998.
 PR 07-OCT-1997; U15753.
 PR 02-OCT-1997; US-061323.
 PR 07-OCT-1996; US-729743.
 PA (UNIV) UNIV JOHNS HOPKINS SCHOOL OF MEDICINE.
 PI Beachy PA, Porter JA;
 DR WPI: 98-399053/34.
 PT New hedgehog-derived poly:peptide(s) - used to develop products for
 PT modulating proliferation or differentiation of neuronal cells,
 PT cholesterol biosynthesis or transport or expression of pituitary
 PT gland gene(s)
 PS Claim 13: Pages 150-151; 210pp; English.
 CC This represents a zebrafish sonic hedgehog (shh) protein sequence. The
 CC invention provides methods and compounds for modulating proliferation or
 CC differentiation of neuronal cells, cholesterol biosynthesis or transport
 CC or expression of pituitary gland genes. The method for affecting
 CC cholesterol biosynthesis or transport in a cell comprises contacting a
 CC cell with a compound that affects hedgehog, thereby affecting cholesterol
 CC biosynthesis or transport. The methods for inhibiting the neural inducing
 CC activity of a hedgehog polypeptide in cells, and for inducing pituitary
 CC gland gene expression utilizes sequences selected from a zebrafish
 CC tiggly-winkle hedgehog (twbn) protein (W61485), a zebrafish sonic hedgehog
 CC (shh) protein (W61486), a chicken shh protein (W61487) or a mouse shh
 CC protein (W61488). The products and methods provide for compounds which
 CC can affect hedgehog activity. They can be used for treating disorders
 CC which arise from neuronal degeneration or abnormal function. They can
 CC also be used as nerve-sparing agents or in restoring or promoting
 CC appropriate patterning during the healing of major limb trauma. They can
 CC also be used for treating Smith-Lemli-Opitz syndrome. The products can
 CC be used for detection and diagnosis.
 SO Sequence 418 AA:

Query Match 58.6%; Score 1274; DB 1; Length 418;
 Best Local Similarity 61.2%; Pred. No. 1e-121;
 Matches 254; Conservative 46; Mismatches 101; Indels 14; Gaps 4;

5 RLRLPHCLVLLLLVYPAAMGCGPRGVGRRPRKLVPLAYKQFSPVNPVKTKGAS 64
 DB 2 RLRLVLLVSLTSLVY-SGLACGPRGYG-RRHPRKLTPLAYKQFIPVNAKTIAGS 59
 QY 65 GRYEKGARSSEKELTPNYNPDIIFKDEENTGADRLMTORCKDRKLSLAISVMNOMPQ 124
 DB 60 GRYEKGITRNSERKELTPNYNPDIIFKDEENTGADRLMTORCKDRKLSLAISVMNOMPQ 119
 QY 125 VKLRTEGMDSDGHSESLHYEGRAVDITTSDDRNKYGILALAYEAGFDWYYSKA 184
 DB 120 VKLRTEGMDSDGHSESLHYEGRAVDITTSDDRNKYGILALAYEAGFDWYYSKA 179
 QY 185 HVHGSVSEHSAAKTGGCPAGAOVRLESGARVALSAVRPGDVLAMGEDGSTFSDVLI 244
 DB 180 HIRCSVKAENSVAAKSGGCPGSAVLSIQDGGQKAVKDLNKGDDVLAADSAGNLFVSDFI 239
 QY 245 IFDREPHRLAFQVITOPPRRLATLPAHLLETA-DNHTEPARFRATFASHVQGVVLVAG 304
 DB 240 METDSDSTRRVFVIEQPEVEKITLTAHLFLVLDNSTEDLTMTAAVASSVBAQKV 299
 QY 305 LVAGVPG-LQPARVAANS-THVALGAYAPLTGKGTLVEDVVASCPAAVADHHLAQLAFWP 363
 DB 300 MVVDSDGOLKSVIVQRTYTEORGSFAPVTAHGTVYVDRIIASCYAVIEDGGLHLAFAP 359
 QY 364 LRLFL-----SLAMGSMTPGEGVHHYPOLLYLGRLLLEBSFHLGMS 407

Db 360 ARLYTVSSFLPONSRRSNATLQEGVHWYSRLLYOMGTWLDLSNMLHPLGMS 414

RESULT 15

R77341

ID R77341 standard; Protein; 475 AA.

AC R77341.

DT 14-MAR-1996 (first entry)

DE Human sonic hedgehog protein.

KW Human; sonic hedgehog protein; probe; primer; diagnostic;

OS nervous system disorder; gene therapy; antibody.

FT Homo sapiens.

FT Key Location/Qualifiers

FT peptide 24..29

FT misc_difference 463 /note= "conserved sequence (R77349)"

FT W09518856-A1. /note= "unspecified amino acid"

PD 13-JUL-1995.

PF 30-DEC-1994; U14992.

PR 30-DEC-1993; US-176427.

PR 14-DEC-1994; US-356060.

PA (HARD) HARVARD COLLEGE.

PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.

PI Ingham PW, McMahon AP, Tabin CJ;

DR WPI; 95-255060/33.

DR N-PSDB; Q91639.

PT Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful

PT to treat degenerative nervous system disorder(s) and in gene

PT therapy.

PS Claim 17; Page 143-45; 210pp; English.

CC The sequence represents a human sonic hedgehog protein, homologous

CC to a Drosophila hedgehog protein (R77337), and is encoded by a cDNA

CC isolated from a human fetal lung cDNA library. Probes and primers

CC derived from the sonic hedgehog gene may be used as diagnostic

CC agents for neuromuscular, autonomic or central nervous system

CC disorders, and the gene may also be used in gene therapy.

CC Antibodies generated from the protein may be used as therapeutic or

CC research reagents.

CC Sequence 475 AA:

Query Match 58.4%; Score 1271; DB 1; Length 475;

Best Local Similarity 57.7%; Pred. No. 2,5e-121;

Matches 267; Conservative 45; Mismatches 85; Indels 66; Gaps 10;

QY 13 CTVLLLT--LVYPAMGCCPGRVVGSRRRPRKLVPLAYKQSPNPPEKTIGASGRYEKG 70

DB 7 CLLLVVSSLLVCSGLACGPRGFG-KRRHPKRLPLAYKQFIPNVAEKTIGASGRYEKG 65

QY 71 IARSSERFELTPNYPNDIIFKDEENTGADRLMTCRCKDLNLSAISVNNOMPVGLRYT 130

DB 66 ISRNSRFKELTPNYPNDIIFKDEENTGADRLMTCRCKDLNLSAISVNNOMPVGLRYT 125

QY 131 EGMDEDGHHSESLAYEGRAVDITTSDRDNKYGILLARLAVEAGFDWYVESKAHVCSV 190

DB 126 EGMDEDGHHSESLAYEGRAVDITTSDRDNKYGILLARLAVEAGFDWYVESKAHVCSV 185

QY 191 KSEHSAATACTGCGFPAGAOYRLSESGARVALSAVRPDVYLAMGEDSPFTSDVLIFLDR 250

DB 186 KAENSVAAKSGCGFPGSATVHLEQGTKLKVDLSPGDRVLAADDQRLYSDFLTFLDRD 245

QY 251 PHRLAFQVIEITQDPPRLATLPRAHLFTA--DNHT-EPARF-----RATF 294

DB 246 DGAKKVFYIETREPPRLITRAHLITVAPRNDSATGEPEASSGSGPSSGALGPRLAF 305

QY 295 ASHVPGQYVLA---GVPGLOPARVAVS-THVALGAYAPLTKHGLTVEDVVA SCFA 349

DB 306 ASRVRRGQRYVVAERDGRRLPAVHSTLSSEAAGAYAPLTAQGTILINRVLASCYA 365

QY 350 AVADHLLAQLATPPLRFLHSL-----AWGSWTRG 378

DB 366 VIEEHSMARRAFPRILAHLLAALAPARTDRGDSGGGDRGGGGRVALTAPGAADAPG 425

QY 379 ---EGVHWYPQLLYRLGRLLLEEGSFHPLGMS-----GAG 410

DB 426 AGATAGIHWYSQLTQIGTWLDSEALHPLGMAVKSXSXRGAG 468

Search completed: June 5, 2000, 08:16:18
Job time: 2583 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2000, 07:55:36 ; Search time 45.88 seconds
(without alignments)
129.332 Million cell updates/sec

Title: US-08-900-220-16

Perfect score: 2175
Sequence: 1 MSPARLRPRHLFCVLVLLLL.....GRLLLEGGSFHPLGMSGAGS 411

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 145308 seqs, 14437401 residues

Total number of hits satisfying chosen parameters: 145308

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/PCITUS.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/backfilest1.pep:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1686	77.5	336	1	US-08-176-427B-6
2	1686	77.5	336	2	US-08-356-060A-10
3	1646	75.7	313	2	US-08-356-060A-14
4	1279.5	58.8	437	4	PCT-US95-15463-20
5	1279.5	58.8	437	4	PCT-US95-15923-20
6	1276.5	58.7	437	1	US-08-176-427B-8
7	1276.5	58.7	437	2	US-08-356-060A-11
8	1274	58.6	418	4	PCT-US95-15463-18
9	1274	58.6	418	4	PCT-US95-15923-18
10	1271	58.4	475	2	US-08-356-060A-13
11	1270.5	58.4	425	1	US-08-176-427B-2
12	1270.5	58.4	425	2	US-08-356-060A-8
13	1268.5	58.3	462	1	US-08-748-591-4
14	1268.5	58.3	462	1	US-08-748-591-9
15	1246.5	57.3	437	1	PCT-US95-02315-2
16	1227	56.4	418	1	US-08-176-427B-10
17	1227	56.4	418	2	US-08-356-060A-12
18	1226.5	56.4	416	4	PCT-US95-15463-17
19	1226.5	56.4	416	4	PCT-US95-15923-17
20	1190.5	54.7	425	4	PCT-US95-15463-19
21	1190.5	54.7	425	4	PCT-US95-15923-19
22	1175	54.0	396	1	US-08-176-427B-4
23	1175	54.0	396	2	US-08-356-060A-9
24	893.5	41.1	471	1	US-08-176-427B-11
25	893.5	41.1	471	2	US-08-356-060A-34
26	775.5	35.7	165	2	US-08-356-060A-41
27	610.5	28.1	167	2	US-08-356-060A-41
28	381	17.5	73	1	US-08-176-427B-13
29	381	17.5	73	2	US-08-356-060A-36

30	346	15.9	73	1	US-08-176-427B-12	Sequence 12, Appl
31	346	15.9	73	2	US-08-356-060A-35	Sequence 35, Appl
32	315	14.5	64	1	US-08-176-427B-15	Sequence 15, Appl
33	315	14.5	64	2	US-08-356-060A-15	Sequence 15, Appl
34	314	14.4	64	1	US-08-176-427B-17	Sequence 17, Appl
35	314	14.4	64	2	US-08-356-060A-17	Sequence 17, Appl
36	300.5	13.8	65	1	US-08-176-427B-16	Sequence 16, Appl
37	300.5	13.8	65	2	US-08-356-060A-16	Sequence 16, Appl
38	263	12.1	64	1	US-08-176-427B-14	Sequence 14, Appl
39	263	12.1	64	2	US-08-356-060A-37	Sequence 37, Appl
40	167	7.7	30	1	US-08-748-591-3	Sequence 3, Appl
41	164	7.5	30	1	US-08-748-591-2	Sequence 2, Appl
42	147	6.8	30	1	US-08-748-591-1	Sequence 1, Appl
43	145	6.7	27	1	US-08-748-591-8	Sequence 8, Appl
44	139	6.4	27	1	US-08-748-591-7	Sequence 7, Appl
45	128	5.9	27	1	US-08-748-591-6	Sequence 6, Appl

ALIGNMENTS

```
RESULT 1
US-08-176-427B-6
; Sequence 6, Application US/08176427B
; Patent No. 5789543
;
; GENERAL INFORMATION:
; APPLICANT: Ingham, Phillip W.
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Tablin, Clifford J.
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
; NUMBER OF INVENTION: Proteins and Uses Related Thereto
; CORRESPONDENCE ADDRESSES: 33
; ADDRESS: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/176,427B
; FILING DATE: 30-DEC-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMI-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 336 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-176-427B-6
```

Query Match 77.5% Score 1686; DB 1; Length 336;
Best Local Similarity 94.9% Pred. No. 2,1e+168;
Matches 319; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

```
OY 76 ERFKELTPNYNDITFKKEENTGADRLMTORCKRLNSLAISVNMQNGVKLRVTEGMD 135
DB 1 ERFKELTPNYNDITFKKEENTGADRLMTORCKRLNSLAISVNMQNGVKLRVTEGMD 60
OY 136 DGHSEESIHYEGRAVDITTSDDRNKYGLLARLAVEAGFDWVYYSRAHWYCSVKSSEHS 195
```

Db 61 DGHSEESLHYEGRADITTSDBDRNKGILLARLAVAGFDMVYTESKAHVCSVSEHS 120
QY 196 AAKTGCCPPAGAOVRLSESGARVALSAVRPGDRLVAMGEDGPTFSDVLIETDREPHRLR 255
Db 121 AAKTGCCPPAGAOVRLSESGARVALSAVRPGDRLVAMGEDGPTFSDVLIETDREPHRLR 180
QY 256 APOVITODPPRRLLATPRLHLLFTADNHTPEAPRATPASHVQPGQYLVAGVPGLOPA 315
Db 181 APOVITODPPRRLLATPRLHLLFTADNHTPEAPRATPASHVQPGQYLVAGVPGLOPA 240
QY 316 RVAAYSTHVALGAYAPLTKHGTLVVEDVYVASCFAAVADHHLAQLAFWPLRFLPSLAWGSM 375
Db 241 RVAAYSTHVALGAYAPLTKHGTLVVEDVYVASCFAAVADHHLAQLAFWPLRFLPSLAWGSM 300
QY 376 TPGEVHWYPMOLLYRLGRLLLEEGSFHPLGMSGAGS 411
Db 301 TPGEVHWYPMOLLYRLGRLLLEEGSFHPLGMSGAGS 336

RESULT 2

US-08-356-060A-10
Sequence 10, Application US/08356060A
Patent No. 5844079
GENERAL INFORMATION:
APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tablin, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
TITLE OF INVENTION: Proteins and Uses Related Thereto
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESSES:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,060A
FILING DATE: 14-DEC-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/176,427
FILING DATE: 30-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMI-006CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 336 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-356-060A-10

Query Match 77.5%; Score 1686; DB 2; Length 336;
Best Local Similarity 94.9%; Pred. No. 2.1e-168;
Matches 319; Conservative 9; Mismatches 8; Indels 0; Gaps 0;
QY 76 EREKELTPYNPDIIFKDEENTGADRLMTORCKDRNLNSLAISVNMOPGVKLVTEGMD 135
Db 1 EREKELTPYNPDIIFKDEENTGADRLMTORCKDRNLNSLAISVNMOPGVKLVTEGMD 60

QY 136 DGHSEESLHYEGRADITTSDBDRNKGILLARLAVAGFDMVYTESKAHVCSVSEHS 195
Db 61 DGHSEESLHYEGRADITTSDBDRNKGILLARLAVAGFDMVYTESKAHVCSVSEHS 120
QY 196 AAKTGCCPPAGAOVRLSESGARVALSAVRPGDRLVAMGEDGPTFSDVLIETDREPHRLR 255
Db 121 AAKTGCCPPAGAOVRLSESGARVALSAVRPGDRLVAMGEDGPTFSDVLIETDREPHRLR 180
QY 256 APOVITODPPRRLLATPRLHLLFTADNHTPEAPRATPASHVQPGQYLVAGVPGLOPA 315
Db 181 APOVITODPPRRLLATPRLHLLFTADNHTPEAPRATPASHVQPGQYLVAGVPGLOPA 240
QY 316 RVAAYSTHVALGAYAPLTKHGTLVVEDVYVASCFAAVADHHLAQLAFWPLRFLPSLAWGSM 375
Db 241 RVAAYSTHVALGAYAPLTKHGTLVVEDVYVASCFAAVADHHLAQLAFWPLRFLPSLAWGSM 300
QY 376 TPGEVHWYPMOLLYRLGRLLLEEGSFHPLGMSGAGS 411
Db 301 TPGEVHWYPMOLLYRLGRLLLEEGSFHPLGMSGAGS 336

RESULT 3

US-08-356-060A-14
Sequence 14, Application US/08356060A
Patent No. 5844079
GENERAL INFORMATION:
APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tablin, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
TITLE OF INVENTION: Proteins and Uses Related Thereto
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESSES:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,060A
FILING DATE: 14-DEC-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/176,427
FILING DATE: 30-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMI-006CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 313 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-356-060A-14

Query Match 75.7%; Score 1646; DB 2; Length 313;
Best Local Similarity 100.0%; Pred. No. 2.8e-164;
Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 101 RMTORCKDRNLNSLAISVNMOPGVKLVTEGMDGHHSESLHYEGRADITTSDBDR 160

```

Db 2 RLTQRCCKDNLNLSIAVNMQGVKRLATYEGWDEDDHSEESLHYGRVADITTSRDR 61
QY 161 NKYGLLARLAVEAGFDWVYVESKAHVCSVSEHSAAKTGCGCPAGAOVRLSEGARVAL 220
Db 62 NKYGLLARLAVEAGFDWVYVESKAHVCSVSEHSAAKTGCGCPAGAOVRLSEGARVAL 121
QY 221 SAVRPGDRLVAMGEDGSPITSVDLIFLDRPHRLRAFOVLETODPPRLALTPAHLFT 280
Db 122 SAVRPGDRLVAMGEDGSPITSVDLIFLDRPHRLRAFOVLETODPPRLALTPAHLFT 181
QY 281 DNHTEPARFRATFASHVQGOVYLVAGVPGLOPARVAVSTHVALGAVAPLTKHGTLLV 340
Db 182 DNHTEPARFRATFASHVQGOVYLVAGVPGLOPARVAVSTHVALGAVAPLTKHGTLLV 241
QY 341 EDVYASCFEAAVADHHLAQLAFWPLRLPHSLAMGSWTPGEGVHWYPLLRLGRLLLEGS 400
Db 242 EDVYASCFEAAVADHHLAQLAFWPLRLPHSLAMGSWTPGEGVHWYPLLRLGRLLLEGS 301
QY 401 FPHLGMGAGS 411
Db 302 FPHLGMGAGS 312

```

RESULT 4
PCT-US95-15463-20
Sequence 20, Application PC/TUS9515463

```

GENERAL INFORMATION:
APPLICANT: The Johns Hopkins University School of Medicine
TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: U.S.A.
ZIP: 92037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/15463
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Halle, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/080M01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-15463-20

```

Query Match 58.8%; Score 1279.5; DB 4; Length 437;
Best Local Similarity 61.1%; Pred. No. 1.2e-125;
Matches 262; Conservative 39; Mismatches 91; Indels 37; Gaps 8;

```

QY 13 CLVLL---LLVPAAMCGRGVGRRRPRKLVPLAYQFSPNPEKTLGASGRYEG 69
Db 8 CFVLTLASSLLVCPGL-ACGPGRGFG-KRRHPKRLTPLAYKQFIPNVAEKTILGASGRYEG 65
QY 70 KIASSERFKELTPNNDIIFKDEENTGADRLTORCKDNLNLSIAVNMQGVKRLV 129

```

```

Db 66 KIRNSERFKELTPTNYPNDIIFKDEENTGADRLTORCKDNLNLSIAVNMQGVKRLV 125
QY 130 TEGWDEDDHSEESLHYGRVADITTSRDRNKYGLLARLAVEAGFDWVYVESKAHVCS 189
Db 126 TEGWDEDDHSEESLHYGRVADITTSRDRNKYGLLARLAVEAGFDWVYVESKAHVCS 185
QY 190 VKSEHSAAKTGCGCPAGAOVRLSEGARVALSAVRPDRVLAMGEDGSPITSVDLIFLDR 249
Db 186 VKSEHSAAKTGCGCPAGAOVRLSEGARVALSAVRPDRVLAMGEDGSPITSVDLIFLDR 245
QY 250 EPHRLRAFOVLETODPPRLALTPAHLFTA-DNHTEPARFRATFASHVQGOVYLVAG 308
Db 246 DEGAKRIFYIETLPERRLTLRAHLFYAPRNDGSPITGPASRRYRPGQRYVVA 305
QY 309 VPG---LQPARVAAS-THVALGAVAPLTKHGTLLVEDVYASCFEAAVADHHLAQLAFWP 363
Db 306 ERGDRLRLPAAVHVSYTLREEGAVYAPLTAHGIILINVLASGVAYIEEHSWHAHFAF 365
QY 364 LRLFHSL-----AMGSWTPGEGVHWYPLLRLGRLLLEE 398
Db 366 FRLAHALLAALAPARDTGGGGGSIIPAQSAATEARGA-EPTAGIHWYSQLLYHIGTWLDS 424
QY 399 GSFHPLGMS 407
Db 425 ETMHPGLMA 433

```

RESULT 5
PCT-US95-15923-20
Sequence 20, Application PC/TUS9515923

```

GENERAL INFORMATION:
APPLICANT: The Johns Hopkins University School of Medicine, et al.
TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: U.S.A.
ZIP: 92037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/15923
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Halle, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/043M01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-15923-20

```

Query Match 58.8%; Score 1279.5; DB 4; Length 437;
Best Local Similarity 61.1%; Pred. No. 1.2e-125;
Matches 262; Conservative 39; Mismatches 91; Indels 37; Gaps 8;

```

QY 13 CLVLL---LLVPAAMCGRGVGRRRPRKLVPLAYQFSPNPEKTLGASGRYEG 69
Db 8 CFVLTLASSLLVCPGL-ACGPGRGFG-KRRHPKRLTPLAYKQFIPNVAEKTILGASGRYEG 65

```

QY	70	KIASSSEFEKLETPNYNDITFKDEENMGADRLTOKCKDRLNSLAISVNMOMGYKLRY	129
Dd	66	KITRNSEFFELIPNYNDILFKDEENTGADRLTOKCKDLNLALASVMOMGYKLRY	1255
QY	130	TGEWDEDGHNSESLHYEGRAVDITTSDDRONKXGLLARLAVEGFPMWYESKAHVCS	189
Dd	126	TGEWDEDGHNSESLHYEGRAVDITTSDDRSKXGMLARLAVEGFPMWYESAHVCS	185
QY	190	VKSEHSAAKTGCGCEPPAGAOVRLSSGARVALSAYRPGEDRVLANMGEDSGPFFSDVLFLDR	249
Dd	186	VKAENSVAANKSGCGCFPGSATVHLBOGGKTKLKIDLRPDRLAADOGRLYSOFLEFLDR	245
QY	250	EPRHLRAOVLETODPPRRLLTFBAHILETA-DNHTEPARAFRATFASHPOGGUYLYAG	308
Dd	246	DEGAKKYFTVETLEPRERLLTHAHLLEFAPHNDSCGTPGESPALSERYRPQGRYIVA	305
QY	309	VPG-----LOPARVAVS-THVALGAVAPLTNGTLVVEDVASCSFAVADHHLAQLFMP	363
Dd	306	ERGDDRLLPAVAHSVTLREBEAGAVAPLTANHGTILLNRVLASCAYITEHSMAHFAP	3655
QY	364	LRLHSL-----ANGSMTPGECVHWYQOLLIRLGRLLEE	398
Dd	366	FRLHALALLAAPARTDGGGGSIPAASOASTEARGA-EPTAGIHWSQLLIHGWTLLDS	424
QY	399	GSHPLIGMS 407	
Dd	425	ETMHPLGMA 433	

RESULT 6
 US-08-176-427b-8
 Sequence 8, Application US/08176427B
 Patent No. 5789543
 GENERAL INFORMATION:
 APPLICANT: Ingham, Phillip W.
 APPLICANT: McMahon, Andrew P.
 APPLICANT: Tablin, Clifford J.
 TITLE OF INVENTION: Vertebrate Embryonic Pattern-inducing
 TITLE OF INVENTION: Proteins and Uses Related Thereto
 NUMBER OF SEQUENCES: 33
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: LAHIVE & COCKFIELD
 STREET: 60 State Street
 City: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: ASCII(text)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/176,427B
 FILING DATE: 30-DEC-1993
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Vincent, Matthew P.
 REGISTRATION NUMBER: 36,709
 REFERENCE/DOCKET NUMBER: HMI-006
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 227-7400
 TELEFAX: (617) 227-5941
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 437 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-176-427b-8

Query Match	58.7%;	Score 1276.5;	DB 1;	Length 437;
Best Local Similarity	60.8%;	Pred. No. 2.4e-125;		
Matches 261;	Conservative 40;	Mismatches 91;	Indels 37;	Gaps 8;

OY	13	CLVLL-----LLVVPAAMCGGFRVVGSSRRPRKLVPLATYKOFSPNDEKTLGASRGREG	69
Dd	8	CFLVYLASSLVCBEL-ACSGFGRGG-KRRRPKILPLAYQOIFPNAVEKTLGASRGREG	65
OY	70	KIARSSERFKELTPNYNDIIKFKEENTGAURLMTOQRCKRLNSIAISVMNQDPVKLRV	129
Dd	66	KITRSEFFKELTPNYNDIIFKDEENTGAURLMTOQRCKRLNSIAISVMNQDPVKLRV	122
OY	130	TCEMEDCHHEESIHYBGRVADTTSPDRNKXGLLARLAVEGFPMVYYEASHNCYS	188
Dd	126	TEGWEDDHHEESITHGRAVDITTSRDKSKGMLARLAVEGFDMVYESKAHNCYS	187
OY	190	VKSEHSAAAKGGCFEPAGAOVRLESGAVALSAVRPGDVLAMEDEGSPTPSDLIEFLR	245
Dd	186	VKAENSVAAKSGGCPGSAVTHLBOGGKTKLVKDLRPGDRVLAADOGRLYSDELTFIDR	242
OY	250	EPRHLRAQVLETODPPRRLLTAHLLFTA-DNHTEPAAFKRTFPASHNOFGQYVLVAG	307
Dd	246	DEGAKRVYELETEPRBRLLLTAAHLLFTVAHPNDSGTPEPSALLFSVRNPGRGVYVA	303
OY	309	VPG-----IQPARVAVS-TVALGAVAPLTKHGTLYVEDVASCSFAAADHNLQLAEWP	367
Dd	306	ERGGRRLLPRAVSVTYLREEDAQAYPDLTHGILLIRVLASCYAVILEHSMHRRFAP	366
OY	364	LRLHSL-----ANGSWTPGEGVHWYDQLLYRLGRLLLEE	397
Dd	366	FRLAHALLAALAPARTDGGGGSIPAAQSATEARGA-EPTAGIMWYSOLLYHTIGTWLDS	424
OY	399	GSHFLIGMS 407	
Dd	425	ETMHPLGMA 433	

RESULT 7
US-08-356-060A-11
Sequence 11, Application US/08356060A
Patent No. 5844079
GENERAL INFORMATION:
APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tabin, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
TITLE OF INVENTION: Proteins and Uses Related Thereto
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
City: Boston
STATE: MA
COUNTRY: USA
Zip: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII/text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,060A
FILING DATE: 14-DEC-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/176,427
FILING DATE: 30-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMT-006CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400


```

: TELEFAX: (617) 227-5941
: INFORMATION FOR SEQ ID NO: 11
:
: SEQUENCE CHARACTERISTICS:
:
: LENGTH: 437 amino acids
:
: TYPE: amino acid
:
: TOPOLOGY: linear
:
: MOLECULE TYPE: protein
:
: US-08-356-060A-11

```

Query Match	58.7%	Score 1276.5	DB-2	Length 437
Best Local Similarity	60.8%	Pred No. 2.4e-125		
Matches 261	Conservative 40	Mismatches 91	Indels 37	Gaps 8

```

OY      13  CLVLL- --LLVWAAMCGGGRVVGSGRRPRKVLVLAKEFSPNVEITLGSRGYEG 69
Db      8  CFVYLLASLLVCGGL-ACGGRGFG- KRHRPKVLPLATKQFLPNAEETLGSRGYEG 65

OY      70  KIANSSERKELLTNNYNDIIFKDEENTGADRLMTORCKRLNSLAISVNMOPGVKRV 139
Db      66  KITRNSSEFKELLTNNYNDIIFKDEENTGADRLMTORCKRLNSLAISVNMOPGVKRV 125

OY      130  TEGWDEDSHHEESLHYEGRAVDITTSROBNKXGLLAFLAVEGPMVYEGSAHVHCS 139
Db      126  TEGWDEDSHHEESLHYEGRAVDITTSROBNKXGLLAFLAVEGPMVYEGSAHVHCS 185

OY      190  VKSEHSAAKGCGCFPGAGQVRLTESGARVAISAAPCPDRVLAMEGDSPTFSVDLFIPLDR 24.9
Db      186  VKAESVSAKSGCGCFPGAGVTHLEGGGKVLKDLRPGDRVLAAODGGRLLYSDFLELDR 24.5

OY      250  EPHRLRAQVLETODPPRRLLTAHLLFTV- DNHTEPARFRATFSHVOPOGYVLVAG 308
Db      246  DEGAKEVYLETLEPERBRLLLTAHLLFTVAPNHDSCGTPGSAISRVRPGORYVVA 305

OY      309  VPG----LQPARVAVS-THVALATAYPLTNGTLVEDVVASGFAVAADHLLAQLAFWP 363
Db      306  ERGDRRLPLPAVHSVTLREEEAQAYPALTNGITLLNRVLASCIAYIEESHVNRHAFAP 365

OY      364  LRLFLST-----ANGSWTPGEGVMYDQLLYRLGRLLLEE 398
Db      366  FRLAALLAALAPARTDGGGGSIPAAQASATEARGA-EPTGIMVYSQLLYHTIGTWLDS 424

OY      399  GSFHLGMS 407
Db      425  ETMRPLGMA 433

```

RESULT 8
PCT-US95-15463-18
: Sequence 18, Application PC/TUS9515463
: GENERAL INFORMATION:
: APPLICANT: The Johns Hopkins University School of Medicine
: TITLE OF INVENTION: NOVEL HEDEHOG-DERIVED POLYPEPTIDES
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 4225 Executive Square, Suite 1400
: CITY: LA Jolla
: STATE: CA
: COUNTRY: U.S.A.
: ZIP: 92037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/15463
: FILING DATE: 01-DEC-1995
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Haile, Lisa A.
: REGISTRATION NUMBER: 38,347

```

? REFERENCE/DOCKET NUMBER: 07265/080W01
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 619/678-5070
? TELEFAX: 619/678-5099
? INFORMATION FOR SEQ ID NO: 18:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 418 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? PCT-US95-15463-18

```

Query Match:	58.68:	Score 1274:	DB 4:	length 418:
Best Local Similarity:	61.28:	Pred. No. 4.1e125:		
Matches 254:	Conservative 46:	Mismatches 101:	Indels 14:	Gaps 4

```

0Y 5 RLRLPHLCVLLLLLVAPAMGGPGVWVSRRRPKRLVPLAYKOFSPVPEKTLGAS 64
Db 2 RLRLRVLLVSLTLTSLVY-SGLACPGRGYG-RRRHRKKLLPLAYKOFIPVNAEKTIGAS 59
0Y 65 GRYECKIARSSERFELTPNYPNDIIFKDEKENTGADRLMTORCKDRJNSLAISVMONPG 124
Db 60 GRYECKIIRNSERFELTPNYPNDIIFDEKENTGADRLMTORCKDRJNSLAISVMNHPG 119
0Y 125 VKLRTCEWDEGDHHSSESLHYEGRAVDITTSDDRNKLTGLANLAYEAGFDWYIYESKA 188
Db 120 VKLRTCEWDEGDHHSSESLHYEGRADVITTSDDKRSKTYGLSLAYEAGFDWYIYESKA 179
0Y 185 HVHCVKSEHSAAKTGCGCFPAGOVRLSEGARVALSVPRGSDVILAMGEDGSPTEFDVL 244
Db 180 HVHCVKSEHSAAKTGCGCFPAGSLVSLDGDGCAVNDLNDGDDVYLALDASAGNLVFSFTI 233
0Y 245 IFLDEPRLRAFOYIETODPPRLALPAHLETTADNHTPEPARFRATEFAHVOFGOVY 304
Db 240 METDRDSTTRAFVYIETQEPEVKITTLAHLLETVLONSTEDTLMTAAVSAVYRAGKV 299
0Y 305 LVAGVPG-LQPARVAANSTHVALGAYAPLTHGCLVVEDVVAVSCFAAVADHHLAQLAFWP 366
Db 300 MVVDDSGOLKSVIYORITTEBORSFPAVTHAGTIYVDRILASCAVYIEDGAGHLHAFAP 358
0Y 364 LRLFH-----SLAMGSWTPGGEVWHYPOLLYRLGLRTHLEDSFHLGLGS 407
Db 360 ARLIYVSSSELPONSSRSRNTALQOEVBWHTSKLLTOMGWTLDSSNMHLGLGS 414

```

RESULT 9
 PCT-US95-15923-18
 ; Sequence 18, Application PC/TUS9515923
 ; GENERAL INFORMATION:
 APPLICANT: The Johns Hopkins University School of Medicine, et al
 TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 4225 Executive Square, Suite 1400
 CITY: La Jolla
 STATE: CA
 COUNTRY: U.S.A.
 ZIP: 92037
 ;
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ;
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/15923
 ; FILING DATE: 04-DEC-1995
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Haile, Lisa A.
 ; REGISTRATION NUMBER: 38,347
 ; REFERENCE/DOCKET NUMBER: 07265/043WO1

```

: TELECOMMUNICATION INFORMATION
:
: TELEPHONE: 619/678-5070
:
: TELEFAX: 619/678-5099
:
: INFORMATION FOR SEQ ID NO: 18:
:
: SEQUENCE CHARACTERISTICS:
:
: LENGTH: 418 amino acids
:
: TYPE: amino acid
:
: TOPOLOGY: linear
:
: - MOLECULE TYPE: protein
:
PCT-US95-15923-18

```

Query Match	58.6%;	Score 1274;	DB 4;	Length 418;
Best Local Similarity	61.2%;	Pred No. 4.1e-125;		
Matches 254;	Conservative 46;	Mismatches 101;	Indels 14;	Gaps 4

[illegible]

RESULT 10
 US-08-356-060A-13
 Sequence 13, Application US/08356060A
 Patent No. 5844079
 GENERAL INFORMATION:
 APPLICANT: Ingham, Phillip W.
 APPLICANT: McMahon, Andrew P.
 APPLICANT: Tablin, Clifford J.
 TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
 TITLE OF INVENTION: Proteins and Uses Related Thereto
 NUMBER OF SEQUENCES: 47
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: LAHAYE & COCKFIELD
 STREET: 60 State Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: ASCII(text)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/356,060A
 FILING DATE: 14-DEC-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:

```

? APPLICATION NUMBER: US 08/176,427
? FILING DATE: 30-DEC-1993
? ATTORNEY/AGENT INFORMATION:
? NAME: Vincent, Matthew P.
? REGISTRATION NUMBER: 36,709
? REFERENCE/DOCKET NUMBER: HWI-006CPE
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (617) 227-7400
? TELEFAX: (617) 227-5941
? INFORMATION FOR SEQ ID NO: 13:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 475 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
US-08-356-060A-13

```

Query Match	58.4%;	Score 1271;	DB 2;	Length 475;
Best Local Similarity	57.7%;	Pred. No. 1e-124;		
Matches 267;	Conservative 45;	Mismatches 85;	Indels 66;	Gaps 10

QY	13	CLVLLLL- LTVPAAMCGCGVRYGSSRRPRPKPLATYKQFSPNVETKLTGASRGYEGK	70
Db	7	CLLVIVSSLLVCGSLACGCGRGCG- KRHRPKLTPLATYKQFIPNVAEKLTGASGRTEGK	65
QY	71	TARSSERFKELTPYNDPIIFKDEENTGADRLMTQRCCKDRLSLAISSVMNQPEVKLRVT	130
Db	66	ISRRSEEFKELTPYNDPIIFKDEENTGADRLMTQRCCKDKLINALAISVMNQPEVKLRVT	122
QY	131	EGMDEGDHSEESIHYEGRAVDITSPDRNRKYGSLRLARLVEAGFDMVYVESKRAHYCSV	199
Db	126	EGMDEGDHSEESIHYEGRAVDITSPDRNRKYGSLRLARLVEAGFDMVYVESKRAHYCSV	188
QY	191	KSESSAAAKTGCGCPAGAOYRLSESGARVALSAVAPGDRVLAAMEGDSPTFSDVLIIFLDRE	250
Db	186	KAENSVAAKSGCGCPGSAATVLEHGCGTKVLDLSPGDRVLAADQGLTLYSDFLTEFDRO	244
QY	251	PHRLRAQVYIETQDPPRRRLATPRLPHLLFTA---DNHT-EPAARF-----RATF	294
Db	246	DGAKKVVYIETREPRERLLTALHLLFVAPHNDSATGEEPEASSGSPGSGALGPALF	303
QY	295	ASHYQPOQYLVLA----GVPGIOPARVAAVS-TVALTGAVAPLTKHGTLVVEDVAVASCPA	348
Db	306	ASRRPRQORYVVAERBGDRRLRPAVHSTVLSSEAAGATAPLTAGSTIILNRVLASCYA	363
QY	350	AVADHLLAQLAFWELRLEHSL-----AWGSMTPG	378
Db	366	VIEHSWAHRAFAFRLAHLLALAPARTDRGDSGGDPRGGCGGRVALTPACADAPG	422
QY	379	-----EGVHWYRQLLYRGRLLLEGGSHPRIGMS-----GAG	410
Db	426	AGATAGLHWYSQDLYQIGTWLNLSEALHPRIGMAVKKSSXSRGAG	468

RESULT 11
US-08-176-427B-2
Sequence 2, Application US/08176427B
Patient No. 5785543
GENERAL INFORMATION:
APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tablin, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
TITLE OF INVENTION: Proteins and Uses Related Thereto
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

```
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/176,427B
FILING DATE: 30-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMI-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 425 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-176-427B-2
```

```
Query Match 58.4%; Score 1270.5; DB 1; Length 425;
Best Local Similarity 63.0%; Pred. No. 9.7e-125;
Matches 257; Conservative 38; Mismatches 96; Indels 17; Gaps 5;
```

```
QY 20 LVVPAAMGCGPGRVGSRPRPKLVPLAYKOFSPVPEKTTGASGRYGGKILARSSEPK 79
DB 19 LVVSGLTGPGRGIG-KRRHKKLTPLAYKOFIPVNAEKTLGASGRYGGKILARSSEPK 77
QY 80 ELTPNNPDIIFKDEBENTGADRLMTORCKDRLNSLAISVNMOPGVKLVTEGMDGDGH 139
DB 78 ELTPNNPDIIFKDEBENTGADRLMTORCKDRLNSLAISVNMOPGVKLVTEGMDGDGH 137
QY 140 SESLSHYEGRADVITTSDDRNKYGILARLAVAGFDWVYVESKAHVCSVSESAAK 199
DB 138 SESLSHYEGRADVITTSDDRNKYGILARLAVAGFDWVYVESKAHVCSVSESAAK 197
QY 200 TGCGPAGAGVPLESGARVALSAVRPGDRLVAMGEDGSPTEFSDVLFLEDRPHRLRAFOY 259
DB 198 SGCGEPGSAIVHLEHGTLVLDLSPGDRLVADADGRLLYSDFLEDRMDSRRLEFY 257
QY 260 IETODPRRLATLPAHLLETTADNHTPEPA-----RFRATFASHVPOQOYVLVAGVPLQ-- 314
DB 258 IETODPRRLATLPAHLLETTADNHTPEPA-----RFRATFASHVPOQOYVLVAGVPLQ-- 317
QY 314 PARVAAYS-THVALGAYADLTGKGLTVEDVYASCFPAVAADHHLAQLAEMPLRLFHSGLAM 372
DB 318 PASVHSVLSREESAGAYADLTGKGLTVEDVYASCFPAVAADHHLAQLAEMPLRLFHSGLAM 377
QY 373 G-----SWTPGEGVHWYPOLLYRLGRLLLEBSFHPGLMSGSGS 411
DB 378 ALCPDGAIPTAATTTTGHWSRLYRIGSWVLDGDLHPGLGKAVAPAS 425
```

```
RESULT 12
US-08-356-060A-8
Sequence 8, Application US/08356060A
Patent No. 5844079
GENERAL INFORMATION:
APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tablin, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
```

```
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,060A
FILING DATE: 14-DEC-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/176,427
FILING DATE: 30-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMI-006CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 425 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-356-060A-8
```

```
Query Match 58.4%; Score 1270.5; DB 2; Length 425;
Best Local Similarity 63.0%; Pred. No. 9.7e-125;
Matches 257; Conservative 38; Mismatches 96; Indels 17; Gaps 5;
```

```
QY 20 LVVPAAMGCGPGRVGSRPRPKLVPLAYKOFSPVPEKTTGASGRYGGKILARSSEPK 79
DB 19 LVVSGLTGPGRGIG-KRRHKKLTPLAYKOFIPVNAEKTLGASGRYGGKILARSSEPK 77
QY 80 ELTPNNPDIIFKDEBENTGADRLMTORCKDRLNSLAISVNMOPGVKLVTEGMDGDGH 139
DB 78 ELTPNNPDIIFKDEBENTGADRLMTORCKDRLNSLAISVNMOPGVKLVTEGMDGDGH 137
QY 140 SESLSHYEGRADVITTSDDRNKYGILARLAVAGFDWVYVESKAHVCSVSESAAK 199
DB 138 SESLSHYEGRADVITTSDDRNKYGILARLAVAGFDWVYVESKAHVCSVSESAAK 197
QY 200 TGCGPAGAGVPLESGARVALSAVRPGDRLVAMGEDGSPTEFSDVLFLEDRPHRLRAFOY 259
DB 198 SGCGEPGSAIVHLEHGTLVLDLSPGDRLVADADGRLLYSDFLEDRMDSRRLEFY 257
QY 260 IETODPRRLATLPAHLLETTADNHTPEPA-----RFRATFASHVPOQOYVLVAGVPLQ-- 314
DB 258 IETODPRRLATLPAHLLETTADNHTPEPA-----RFRATFASHVPOQOYVLVAGVPLQ-- 317
QY 314 PARVAAYS-THVALGAYADLTGKGLTVEDVYASCFPAVAADHHLAQLAEMPLRLFHSGLAM 372
DB 318 PASVHSVLSREESAGAYADLTGKGLTVEDVYASCFPAVAADHHLAQLAEMPLRLFHSGLAM 377
QY 373 G-----SWTPGEGVHWYPOLLYRLGRLLLEBSFHPGLMSGSGS 411
DB 378 ALCPDGAIPTAATTTTGHWSRLYRIGSWVLDGDLHPGLGKAVAPAS 425
```

```
RESULT 13
US-08-748-591-4
Sequence 4, Application US/08748591
Patent No. 5759811
GENERAL INFORMATION:
APPLICANT: Epstein, Eryn
APPLICANT: Hu, Zhilan
APPLICANT: Bonifas, Jeanette
TITLE OF INVENTION: Mutant Human Hedgehog Gene
```


QY 379 ----EGVHWYQQLYRLGRLLESGSFHPLGMS 407
 DB 426 AGATAGIHWYSQLLYQIGTWLDSALHPLGMA 458

RESULT 15

PCT-US95-02315-2

Sequence 2, Application PC/TUS9502315

GENERAL INFORMATION:

APPLICANT: Jessell, Thomas M.

APPLICANT: Dodd, Jane

APPLICANT: Roelink, Henk

APPLICANT: Edlund, Thomas

TITLE OF INVENTION: DNA ENCODING A VERTEBRATE HOMOLOG OF

TITLE OF INVENTION: HEDGEHOG, VHH-1, EXPRESSED BY THE NOTOCHORD, AND USES

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/02315

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: John P. White

REGISTRATION NUMBER: 28,678

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 278-0400

TELEFAX: (212) 391-0525

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 437 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US95-02315-2

Query Match 57.38; Score 1246.5; DB 4; Length 437;
 Best Local Similarity 59.98; Pred. No. 3.3e-122;
 Matches 257; Conservative 40; Mismatches 95; Indels 37; Gaps 8;

QY 13 CLVLL---LLVPRAMCGGRVYVGRRRPRKLVPLAYKQFSNPVEKTLGASGRYEG 69
 DB 8 CFLVALASSLLVCPGL-ACGPGRGFG-KRQHPKRLTPLAYKQFIPVNAEKTIGASGRYEG 65
 QY 70 KIASSSEFKELTPYNYNDIIFKDEENTGADRLMTQRCCKRLNSLAISVMNOMPQVRLRV 129
 DB 66 KITRNSERFKELTPYNYNDIIFKDEENTGADRLMTQRCCKRLNSLAISVMNOMPQVRLRV 125
 QY 130 TEGWDEDEGHSEESLHYEGRAVDITTSDDRNKYGLLARLAVEAGFDWYVESKAHVCS 189
 DB 126 TEGWDEDEGHSEESLHYEGRAVDITTSDDRNKYGLLARLAVEAGFDWYVESKAHVCS 185
 QY 190 VKSEHSAAKGCGCPPAQOYRLGEGARVALSAVPRGDRVLAMGEDGSPTSVDLIFLDR 249
 DB 186 VKAENSVAKSDGCGPSATVHLEGGGTRKLVKDLSPGDRVLAADQGRLLYSDFLFLDR 245
 QY 250 EPHRLRAFQVETODPPRRLALTPLHLFTA-DNHTPEARFRATFASHVQPGQYVLVAG 308
 DB 246 DEGAKKVYVETREPRERLLTAHLFLVAPHNDSGPTPGSPFLFASRVRPGQRYVYA 305

QY 309 VPG----LQPARVAAS-THVALGAYAPLTKHGLVYEDVYASGFAVADHHLAQLAFWP 363
 DB 306 ERGDRRLPLPAVHVSVTLREBAAGAYAPLTADGTLINRVLASCYAVIEHSMVHRAFAFAP 365
 QY 364 LRLFHSL-----AMGSWTGPGVHWYQQLYRLGRLLEE 398
 DB 366 FRLHALIALAPARTDGGGGGSLPAPQSVAEARAGAPPA-GIHWYSQLLYHIGTWLDS 424
 QY 399 GSFHPLGMS 407
 DB 425 ETLHPLGMA 433

Search completed: June 5, 2000, 08:17:15
 Job time: 1299 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2000, 07:57:20 ; Search time 22.84 Seconds

(Without alignments)
1055.027 Million cell updates/sec

Title: US-08-900-220-16

Perfect score: 2175

Sequence: 1 MSPARLRPRHFCVLLVLL.....GRLLLESGFHLGSGAGS 411

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 168808 seqs, 58629743 residues

Total number of hits satisfying chosen parameters: 168808

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1686	77.5	336	2 C49425	Indian hedgehog pr
2	1276.5	58.7	437	2 A49425	Sonic hedgehog pro
3	1274	58.6	418	2 A53193	hedgehog homolog v
4	1270.5	58.4	425	2 A49424	patterning protein
5	1251.5	57.5	444	2 S56765	morphogen xhh prec
6	1246.5	57.3	437	2 B53193	hedgehog homolog v
7	1227	56.4	415	2 A49426	Sonic hedgehog gen
8	1175	54.0	396	2 B49425	Desert hedgehog pr
9	893.5	41.1	471	2 A46400	segment polarity p
10	456	21.0	94	2 G02735	desert hedgehog -
11	174.5	8.0	615	2 T29550	hypothetical prote
12	173.5	8.0	868	2 T22281	hypothetical prote
13	170.5	7.8	1021	2 T23252	hypothetical prote
14	161	7.4	1226	2 T24045	hypothetical prote
15	154.5	6.8	1207	2 T23754	hypothetical prote
16	148.5	6.8	481	2 T27665	hypothetical prote
17	148	6.4	484	2 T34504	hypothetical prote
18	138.5	6.4	629	2 T19563	hypothetical prote
19	123.5	5.7	205	2 T26220	hypothetical prote
20	105	4.8	1997	2 T30874	virginiamycin S sy
21	102	4.7	751	2 T29357	hypothetical prote
22	101.5	4.7	898	2 S74903	DNA topoisomerase
23	99	4.6	1063	2 B70841	probable helz prot
24	98.5	4.5	316	2 T35006	beta-ketocycl-acyl
25	98	4.5	1222	2 G72614	probable reverse g
26	97	4.5	7463	2 T36248	CDA peptide synthe
27	96	4.4	8563	2 T30226	polyketide synthas
28	95.5	4.4	846	2 H70871	hypothetical prote
29	95	4.4	808	2 T04092	phospholipase D (E
30	95	4.4	1616	2 G70668	polyketide synthas

31	95	4.4	2117	2 T36180	CDA peptide synthe
32	94	4.3	482	2 S65760	dihydroilpoamide t
33	94	4.3	747	1 QRECFE	ferriochrome-iron r
34	93.5	4.3	861	2 B49847	nitrate reductase
35	92.5	4.3	127	2 T35336	Probable isomerase
36	92.5	4.3	413	2 B64983	Yela protein - Esc
37	92.5	4.3	638	2 F75547	anthranilate synth
38	92.5	4.3	736	2 T12963	subtilisin homolo
39	91.5	4.2	333	2 S75980	hypothetical prote
40	91	4.2	6260	2 T30228	polyketide synthas
41	91	4.2	10223	2 T30225	polyketide synthas
42	90.5	4.2	1102	2 JH0717	guanylate cyclase
43	90	4.1	382	2 T35709	hypothetical prote
44	90	4.1	609	1 KSA5L1	laccase (EC 1.10.3
45	89.5	4.1	1575	3 T18545	lysobactin synthe

ALIGNMENTS

RESULT 1

C49425

Indian hedgehog protein - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 06-Jan-1995

C:Accession: C49425

R:Rehelard, Y.; Epstein, D.J.; St-Jacques, B.; Shen, L.; Mohler, J.; McMahon, J.A.; M

Cell 75, 1417-1430, 1993

A:Title: Sonic hedgehog, a member of a family of putative signaling molecules, is imp

A:Reference number: A49425; M01D:94094336

A:Accession: C49425

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-336 <ECCH>

A:Cross-references: GB:X76291

Query Match 77.5%; Score 1686; DB 2; Length 336;

Best Local Similarity 94.9%; Pred. No. 1.9e-135;

Matches 319; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY	76	ERFKEIPNPNYNDIIFDEENTGADRIMTORCKRRLSLAISVNNQMPGYKRLRTGEMDE	135
DB	1	ERFKEIPNPNYNDIIFDEENTGADRIMTORCKRRLSLAISVNNQMPGYKRLRTGEMDE	60
QY	136	DGHSEESLHYEGRAVDITTSRDRNRYGLARLAVAGEFDMVYEEKAHVCSEKSEHS	195
DB	61	DGHSEESLHYEGRAVDITTSRDRNRYGLARLAVAGEFDMVYEEKAHVCSEKSEHS	120
QY	196	AAKTGGCFPAQAVRLSGARVALSAVRPGRVLAAGEDSPFFSDVLIPLDEPRRLR	255
DB	121	AAKTGGCFPAQAVRLSGARVALSAVRPGRVLAAGEDSPFFSDVLIPLDEPRRLR	180
QY	256	AFQVETDDPPRLALPFAHLFTADNHTPEPAFRATFASHVPGOVLAAGVGLQPA	315
DB	181	AFQVETDDPPRLALPFAHLFTADNHTPEPAFRATFASHVPGOVLAAGVGLQPA	240
QY	316	RVAAVSTHVALGAYAPLTKHGTIVVEDVVASCFPAVADHHLAQAFWPLRLFLSLANGSW	375
DB	241	RVAAVSTHVALGAYAPLTKHGTIVVEDVVASCFPAVADHHLAQAFWPLRLFLSLANGSW	300
QY	376	TPGEGVHYRPLLRLRGLLLEBSFHLGSGAGS 411	
DB	301	TPGEGVHYRPLLRLRGLLLEBSFHLGSGAGS 336	

RESULT 2

A49425

Sonic hedgehog protein precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Dec-1999

C:Accession: A49425

R:Rehelard, Y.; Epstein, D.J.; St-Jacques, B.; Shen, L.; Mohler, J.; McMahon, J.A.; M

```

Oy 65 GRYEGKIASRSEKELTPNYPNDIIFKDEENTGADRLMTORCKDRNLSTAVNMOPG 124
    ||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 60 GRYEKITRNSRERKELTPNYPNDIIFKDEENTGADRLMTORCKDRNLSTAVNMOPG 119
Oy 125 VKLRTEGMEDEGHHSESLHYEGRAVDITTSRDRNKKYGLARLAVENAGFDMVYESKA 184
    ||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 120 VKLRTEGMEDEGHHSESLHYEGRAVDITTSRDRNKKYGLARLAVENAGFDMVYESKA 179
Oy 185 HVHCVKSEHSAAKTGCFCFPAQVRLSESGARVALSAVRPDRVLAMGEDSPFSDVL 244
    ||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 180 HVHCVKSEHSAAKTGCFCFPAQVRLSESGARVALSAVRPDRVLAMGEDSPFSDVL 239
Oy 245 IFLDSEPRRLRAFOYIETODPPRLALTPAHLTFADNHTPEPAARFATFASHVOPGQYV 304
    ||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 240 METDRDSTRRYFYIETQEPYKTLTVAHLLFVLNDSTEDLHMTAAVSVRAQGV 299
Oy 305 LVAGVPG-LQPARVAAVSTHVALGAYAPLTGTLVEDVYASCAVAADHHLAOLAMP 363
    ||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 300 MVVDGGLKSVIVQRTTEORGSFAPVTAHTIVDRIILASCAVIEDQGLARLAP 359
Oy 364 LRLFFH--SLAMGWTGCGVHWYPOLLY-----RLGRLLEEGSFHPLGMS 407
    ||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 360 ARLYYVSFLSPKTPANG----PMRLYNRRSGTGTGSCHQMGWTLDSNMLHPLGMS 414

```

RESULT 8

B49425

Desert hedgehog protein precursor - mouse

C:Species: Mus musculus (house mouse)

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Dec-1999

C:Accession: B49425

R:Etcheld, Y.; Epstein, D.J.; St-Jacques, B.; Shen, L.; Mohler, J.; McMahon, J.A.; McMa

A:Title: Sonic hedgehog, a member of a family of putative signaling molecules, is impli

A:Reference number: A49425; MUID:9404334

A:Accession: B49425

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-396 <ECH>

A:Cross-references: GB:X76292; NID:9443941; PIDN:CA53924.1; PID:9443942

C:Genetics:

A:Gene: Dhh

C:Superfamily: sonic hedgehog protein

Query Match 54.0%; Score 1175; DB 2; Length 396;

Best Local Similarity 58.5%; Pred. No. 5.5e-92;

Matches 234; Conservative 50; Mismatches 102; Indels 14; Gaps 5;

Oy 3 PARLPRHFCVLVLLVPPAAMCGGFR--VGSRRPRRLVPLAKQSPNPEKTL 61

Db 4 PASLPLFC--CLALLAL-----SAQSGFGRGVGRKRYRKQVLPLKQVPSNPEKTL 57

Oy 62 GASGHEKIASRSEKELTPNYPNDIIFKDEENTGADRLMTORCKDRNLSTAVNM 121

Db 58 GASGAEERVRKGSERFEDLPNYPNDIIFKDEENSGADRLMTORCKDRNLSTAVNM 117

Oy 122 MPGVLRTEGMEDEGHHSESLHYEGRAVDITTSRDRNKKYGLARLAVENAGFDMVYE 181

Db 118 MPGVLRTEGMEDEGHHSESLHYEGRAVDITTSRDRNKKYGLARLAVENAGFDMVYE 177

Oy 182 SKAHVCSKSEHSAAKTGCFCFPAQVRLSESGARVALSAVRPDRVLAMGEDSPFSS 241

Db 178 SRNHVHVSVKADNSLAVAGCGFPGNATVRLSSGKRLRELHRDWDVLADAAQGVYPT 237

Oy 242 DVLTLDEPRRLRAFOYIETODPPRLALTPAHLTFADNHTPEPAARFATFASHVOPG 301

Db 238 PVLLFLDLDLDRRASVFAVETERPRKLLTPWHLVFAARGPAPAGFAVFAARLRAG 297

Oy 302 QYVLAVGPGLOPARVAAVSTHVALGAYAPLTGTLVEDVYASCAVAADHHLAOLAF 361

Db 298 DSVLAPGGDALQPARVAAVSTHVALGAYAPLTGTLVEDVYASCAVAADHHLAOLAF 357

```

Oy 362 WPLRFLSLAMGWTGPG-----EGVHWYPOLLYRLGRLL 396
    ||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 358 APLRLRLAL--GALLPGAVOPTGMHWYSRLYLRLABELM 395

```

RESULT 9

A46400

segment polarity protein hedgehog - fruit fly (Drosophila melanogaster)

N:Alternate names: hh protein

C:Species: Drosophila melanogaster

C>Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 10-Dec-1999

C:Accession: A46400; JN0501; A43480

R:Tabata, T.; Eaton, S.; Kornberg, T.B.

A:Title: The Drosophila hedgehog gene is expressed specifically in posterior comparn

A:Reference number: A46400; MUID:94040725

A:Accession: A46400

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-471 <TAB>

A:Cross-references: GB:S66384; NID:9435848; PID:9435849

A:Experimental source: Oregon-R

A:Note: sequence extracted from NCBI backbone (NCBIN:138996, NCBI:P.138997)

R:Tabata, S.; Michie, T.; Higashijima, S.; Zennou, S.; Ishimaru, S.; Takahashi, F.;

A:Title: Structure and expression of hedgehog, a Drosophila segment-polarity gene req

A:Reference number: JN0501; MUID:93185922

A:Accession: JN0501

A:Molecule type: mRNA

A:Residues: 1-471 <TAB>

A:Cross-references: GB:U05404

A:Note: It is uncertain whether Met-1 or Met-51 is the initiator

A:Note: Intron positions were determined from partial DNA sequence

R:Lee, J.J.; von Kessler, D.P.; Parks, S.; Beachy, P.A.

A:Title: Secretion and localized transcription suggest a role in positional signaling

A:Reference number: A43480; MUID:93008241

A:Accession: A43480

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-471 <LEE>

A:Cross-references: GB:U02793; NID:9157609; PID:9157610

A:Note: sequence extracted from NCBI backbone (NCBI:P.115418)

C:Comment: This protein is required for cell-cell communication.

C:Genetics:

A:Gene: hh

A:Cross-references: Flybase:FBgn0004644

A:Introns: 160/3; 248/1

C:Superfamily: sonic hedgehog protein

C:Keywords: transmembrane protein

F:62-82/Domain: transmembrane #status predicted <TM>

Query Match 41.1%; Score 893.5; DB 2; Length 471;

Best Local Similarity 45.9%; Pred. No. 5.3e-68;

Matches 189; Conservative 61; Mismatches 131; Indels 31; Gaps 7;

Oy 14 LVLLLVLP-----AAMCGGGRVYSGRRPRPKLVPLAKQSPNPEKTLGASGREG 69

Db 67 LVALLLVLPVWFSPASCGRGRLGRHR--ARNLYLPLVYKQTPNISEYTNSSAGLEG 124

Oy 70 KIASSERFKELTPNYPNDIIFKDEENTGADRLMTORCKDRNLSTAVNMOPGVRLV 129

Db 125 VYRDSRFRKFDLVNYPNDIIFRDEEGTADRLMSKCKEKLNLVLAIVNMENFGITLLV 184

Oy 130 TEGMEDEGHHSESLHYEGRAVDITTSRDRNKKYGLARLAVENAGFDMVYESKAHVCS 189

Db 185 TESMEDVYHHQOESLHYEGRAVITATSDRQSKYGLARLAVENAGFDMVSVSRHLYCS 244

Oy 190 VKSHSAAKTGCFCFPAQVRLSESGARVALSAVRPDRVLAMGEDSPFSDVLTLDR 249

Db 245 VKSDSISSHVGCFTPESTALLESQVRKPLGELISIDRVLSMTANQAVSYEILFMDR 304

Qy 250 EPHRLAFQVETEDDPRLALPAHLFLFADNHEPAPRFATRSHPQGVYLVAGV 309
 Db 305 NLEQKQNFVQLHT-DGCAVLTVIPAHVLSVWQPESO--KLTFVFPADRIEKNQVLTADV 360
 Qy 310 P-GIOPAPAAVASTP~~AL~~ALGAYAPLTKEGTLVEDVYASCFAPVADHILAQLAFWPLRLF 367
 Db 361 ETGELRPQRYVVKGVSVKGVAPLTFEGEITIVNSVAACGVAVINQSLAHNGLAPMILL 420
 Qy 368 HSLANGSTPGE-----GVHWYPOLYRLRGLLEEGSFH 402
 Db 421 STLE--AWLPKAEQLHSSPKVYSSAQOQNGIHWYANALYKVDYVLPQSMRH 470

RESULT 10
 G02735
 desert hedgehog - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 17-Jul-1998
 C:Accession: G02735
 R:Drummond, I. A.
 submitted to the EMBL Data Library, June 1996
 A:Reference number: H01643
 A:Accession: G02735
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-94 <DRU>
 A:Cross-references: EMBL:U059748; NID:g1401271; PTD:g1401272
 C:Genetics:
 A:Gene: hDHH

Query Match	21.0%	Score 456	DB 2	Length 94
Best Local Similarity	86.2%	Pred No. 7	6e-32	
Matches	81	Conservative	11	Mismatches 2
				Indels 0
				Gaps 0

QY	11FKDEENTGADRLMTQRCCKDLNLSAISVNNQMPGVKLVTEGMDDEGHSHSESLHVEG	148
DB	111FKDEENSGARLILTEKREKERYNNALAIYMMWVGVVRVTEGMDDEGHSHSESLHVEG	60

```
QY 149 RAVDITTSDRDRNKGGLARLAVEAGFDWYYES 1822
    ||:|||||
Db 61 RALDITTSDRDRNKGGLARLAVEAGFDWYYGS 94
```

```

RESULT 11
T29550
Hypothetical protein ZK377.1 - Caenorhabditis elegans
C.Species: Caenorhabditis elegans
C.Date: 15-Oct-1999 #sequence-revision 15-Oct-1999 #text-change 15-Oct-1999
C.Accession: T29550
R.Nhan, M.; Hawkins, J.
submitted to the EMBL Data Library, February 1997
A.Description: The sequence of C. elegans cosmid ZK377.
A.Reference number: Z20639
A.Accession: T29550
A.Status: preliminary; translated from GB/EMBL/DDbJ
A.Molecule type: DNA
A.Residues: 1-615 <NHA>
A.Cross-references: PTDN:AA52656.1; GSPDB:GN00028; CESP:ZK377.1
A.Experimental source: strain Bristol N2; clone ZK377
C.Genetics:
A.Gene: CESP:ZK377.1
A.Map position: 10
A.Intons: 46/1; 70/3; 157/2; 176/1; 245/2; 314/3; 402/3; 489/2

```

Query Match	8.0%	Score 174.5	DB 2:	Length 615:
Best Local Similarity	24.7%	Pred 24.7%-07:		
Matches	65:	Conservative 52:	Mismatches 11:	Indels 35:
				Gaps 11:
QY	142	ESLHTEGRAVDDITTSDDRD-	NKYGLIARLAVLEGEDWYEE-	KAHVCSVSEHSAAK 199
DB	342	SGSGSYSGSYGPTLDASQYNAY	PAHQCPAYQPAQPAQOPAYQPA	YQAPAYQPAVSARCY 401

```

0Y 200 T-----GG-----CFPAGAVRLESSARVALSVRGDRILAMGEOSPFSVILFD 248
Db 402 SPNLNGLEFGSGMOCFSDMEVELEDGK- IKDKLIGDKVLSN-DEAVETISFVIELH 459
QY 249 REPRLRAFOVLETQDPRRLALTPAHLETLADNHE-----PAARPRATFASHVOPG 302D
Db 460 KRDEIABFNLEIFAN-GHSIKLDNHLIYSDCRTLSDLKLVAAKRVKDDDCIHVTTDS 518
QY 303 YVLVAGVGLDPAVAAVSTHVALGATAPLTKHTLYVEDVVASCFEAAVADHHLAQLAFW 362D
Db 519 NVYI-----KKRVYSKSKIETIGIYSPSTGDIIVRYLASCHSNLAKLSLOQ-TFF 570
QY 363 PL-----RLFSLAMGSWTPEEG 380
Db 571 SLVKRTSSVFHNLMPFKSSTEEG 593

```

RESULT 12
T22281
hypothetical protein F46B3.5 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T22281
R:Ainscough, R.
submitted to the EMBL Data Library, November 1996
A:Reference number: 219541
A:Accession: T22281
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-868 <MWL>
A:Cross-references: EMBL:T21540; PIDB:CAB04405.1; GSPDB:GN00023; CESP:F46B3.5
C:Experimental source: clone F46B3
C:Genetics:
A:Gene: CESP:F46B3.5
A:Map position: 5
A:Introns: 45/3; 200/1; 235/3; 436/3; 561/3; 561/3; 696/3; 742/1; 826/3; 846/3

[illegible]

```

RESULT      13
T23252
hypothetical protein K02E2.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence
C:Accession: T23252
R:LiJyd, C.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19716
A:Accession: T23252
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1021 <WIL>

```

A:Cross-references: EMBL:281560; PIDN:CA04547.1; GSPDB:GN00023; CESP:K02E2.2
A:Experimental source: clone K02E2
C:Genetics:
A:Gene: CESP:K02E2.2
A:Map position: 5
A:Introns: 41/1; 79/3; 131/3; 197/1; 262/1; 297/3; 566/3; 610/3; 735/3; 775/3; 834/3; 88

Query Match 7.8%; Score 170.5; DB 2; Length 1021;
Best Local Similarity 24.3%; Pred. No. 3.3e-06;
Matches 55; Conservative 42; Mismatches 88; Indels 41; Gaps 7;

OY 203 CFPAGAVRLEGARVALSAVRPGDRLVAMGEDSPFSDVLIPLDREPHRLRAF 262
DB 802 CFSRDTWTFPSGKR-RMDEIGDVL-ADLTALFSAITLWIRHREPEVOELEIKT 859
OY 263 ODPRLALTPAHLFTADNTEPA-----ARFR-----ATFASHVOPG 301
DB 860 -DNGKITQLTACHFTATPCRLPSKNSLSTPERYHLLDLPDSETKLAQLKIG 918
OY 302 QYLVAGVPGLOPARVAVSTHVALGAVAPLTKHGTIVEDVVASCPAAVADHHLAQLAF 361
DB 919 ECLLHNGDOFMOKIDSISKTSTGISTPLTENGRIIVNDVLASCYSEVQCNVLTTF 978
OY 362 WPLRLFHSLANGSWTPGEGVHWYPOLYRLGRLLESGSFHPLGMS 407
DB 979 WAFDRLRLN-----IVQYGDLYDEIEL-PTGTS 1007

RESULT 14

T24045

hypothetical protein R08B4.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T24045

R:White, S.

submitted to the EMBL Data Library, November 1995

A:Reference number: Z19834

A:Accession: T24045

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1226 <WILL>

A:Cross-references: EMBL:268008; PIDN:CAA92000.1; GSPDB:GN00028; CESP:R08B4.1

A:Experimental source: clone R08B4

C:Genetics:

A:Gene: CESP:R08B4.1

A:Map position: 10

A:Introns: 53/2; 113/1; 152/3; 204/3; 272/1; 354/1; 389/3; 714/3; 839/3; 877/1; 960/3; 1

Query Match 7.4%; Score 161; DB 2; Length 1226;
Best Local Similarity 25.0%; Pred. No. 2.7e-05;
Matches 43; Conservative 39; Mismatches 80; Indels 10; Gaps 5;

OY 198 KTGCGCFRAGAGVRLSGARVALSAVRPGDRLVAMGEDSPFSDVLIPLDREPHRLRAF 257
DB 1016 AATGACFSLDITWTFPTGKR-RMDQIDIGDYLTLADL-KTYFTPTLWIRHREPEVOEF 1073
OY 258 QYIEODPPRLALTPAHLF---TADNTEPAARF---RATFASHVOPGQYLVAGVP 310
DB 1074 LTIMT-EYKTIKIRITSRHMWYNNKCKSKYQYIKMLPHDEALFASDLEVGCVVLYKIG 1132
OY 311 GLQPARVAVSTHVALGAVAPLTKHGTIVEDVVASCPAAVADHHLAQLAFW 362
DB 1133 KYROCKIEITRSVTRGYSPLTNNGRIIVNDMLASCYSEIQONTLTQTFEF 1184

RESULT 15

T23754

hypothetical protein T05C12.10 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999

C:Accession: T23754; T24513

R:Thomas, K.
submitted to the EMBL Data Library, June 1995
A:Reference number: Z19793

A:Accession: T23754

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1207 <WILL>

A:Cross-references: EMBL:24966; PIDN:CAA90265.1; GSPDB:GN00020; CESP:T05C12.10

A:Experimental source: clone M10

R:Burton, J.
submitted to the EMBL Data Library, October 1995

A:Reference number: Z19901

A:Accession: T24513

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1207 <WILL>

A:Cross-references: EMBL:266500; PIDN:CAA91313.1; GSPDB:GN00020; CESP:T05C12.10

A:Experimental source: clone T05C12

C:Genetics:

A:Gene: CESP:T05C12.10

A:Map position: 2

A:Introns: 31/3; 87/2; 141/3; 180/2; 203/3; 267/1; 776/2; 794/2; 834/2; 1086/3; 1143/

Query Match 7.1%; Score 154.5; DB 2; Length 1207;
Best Local Similarity 28.2%; Pred. No. 9.5e-05;
Matches 51; Conservative 30; Mismatches 83; Indels 17; Gaps 5;

OY 196 AAATGCG---CFPAGAVRLEGARVALSAVRPGDRLVAMGEDSPFSDVLIPLDREP 251
DB 950 AAGAGGGRSNCFSADSLVTVTGOK-RMDELQIDIGDYLVSNGVNLKYEKEMFTHREP 1008
OY 252 HRLRAFQVIEODPPRLALTPAHLF-----FTADNTEPAARFRAATFASHVOPG 301
DB 1009 KTRTFEVLVLYTKS-CRKLSTLGRHLLPVAECQYEQYTMNPDGIDVAMRESKYAEKARKG 1067
OY 302 QYLVAGVPGLOPARVAVSTHVALGAVAPLTKHGTIVEDVVASCPAAVADHHLAQLA 360
DB 1068 ECVLSIDSGEVIADIVRGVMTNVGISPMTEGSLIVDGVLSGSPSHLSSHAKLI 1127
OY 361 F 361
DB 1128 F 1128

Search completed: June 5, 2000, 08:17:52
Job time: 1232 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2000, 08:17:16 ; Search time 14.43 Seconds

(without alignments)
867.429 Million cell updates/sec

Title: us-08-900-220-16

Perfect score: 2175
Sequence: 1 MSPARLRPRHLCVLLILL.....GRLLREGSFHPIGMSGAS 411

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 83857 seqs, 30454973 residues

Total number of hits satisfying chosen parameters: 83857

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : SwissProt_38:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	2162	99.4	411 1	IHH_HUMAN
2	2069	95.1	411 1	IHH_MOUSE
3	1645.5	75.7	408 1	IHH_CHICK
4	1483	68.2	409 1	IHH_XENLA
5	1279.3	58.8	437 1	SHH_MOUSE
6	1274	58.6	418 1	SHH_BRARE
7	1270.5	58.4	425 1	SHH_CHICK
8	1268.5	58.3	462 1	SHH_HUMAN
9	1263	58.1	432 1	SHH_CYNPY
10	1251.5	57.5	444 1	SHH_XENLA
11	1246.5	57.3	437 1	SHH_RAT
12	1226.5	56.4	416 1	TWHH_BRARE
13	1196	55.0	412 1	IHH_BRARE
14	1175	54.0	396 1	DHH_MOUSE
15	1171.5	53.9	396 1	DHH_HUMAN
16	1117	51.4	398 1	DHR2_XENLA
17	1114.5	51.2	396 1	DHR1_XENLA
18	882.5	40.6	471 1	HH_DROME
19	870	40.0	481 1	HH_DROH
20	555	25.5	121 1	SHH_CARAU
21	555	25.5	121 1	SHH_PUNTE
22	553	25.4	121 1	SHH_TANAL
23	552	25.4	121 1	SHH_ANCEH
24	552	25.4	121 1	SHH_RASEL
25	551	25.3	121 1	SHH_RASPA
26	550	25.3	121 1	SHH_DANAA
27	550	25.3	121 1	SHH_DANAE
28	550	25.3	121 1	SHH_DANAT
29	550	25.3	121 1	SHH_DANER
30	550	25.3	121 1	SHH_DANKE
31	550	25.3	121 1	SHH_DANPU
32	550	25.3	121 1	SHH_DEUDE
33	550	25.3	121 1	SHH_DEVMA
34	550	25.3	121 1	SHH_DEVPA

35	548	25.2	121 1	SHH_RASHE	P79864 rasbora het
36	546	25.1	121 1	SHH_PUNCO	P79838 punctus con
37	295	13.6	58 1	IHH_CARAU	P79693 carassius a
38	295	13.6	58 1	IHH_DANAT	O13240 danio aff.
39	295	13.6	58 1	IHH_DANKE	P79711 danio kerri
40	295	13.6	58 1	IHH_DANPU	P79713 danio pulch
41	295	13.6	58 1	IHH_DEUDE	O13243 devario dev
42	295	13.6	58 1	IHH_PUNTE	P79852 punctus tet
43	295	13.6	58 1	IHH_RASEL	P79860 rasbora ele
44	290	13.3	58 1	DHH_BRARE	P79729 brachydanio
45	290	13.3	58 1	DHH_DANKE	P79712 danio kerri

ALIGNMENTS

RESULT 1

IHH_HUMAN STANDARD; PRT; 411 AA.

ID IHH_HUMAN

AC O14623; Q43322;

DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE INDIAN HEDGEHOG PROTEIN PRECURSOR (IHH) (HNG-2).

GN IHH.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;

CC Eutheria; Primates; Catarrhini; Homiidae; Homo.

RN [1]

RP SEQUENCE FROM N.A.

RA Tate G., Endo Y., Mitsuura T.;

RT "Human Indian Hedgehog.";

RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE OF 100-411 FROM N.A.

RC TISSUE-FETAL LUNG;

RX MEDLINE; 96070431.

RA Marigo V., Roberts D.J., Lee S.M.K., Tsukurov O., Levi T.,

RA Gaster J.M., Epstein D.J., Gilbert D.J., Copeland N.G., Seidman C.E.,

RA Jenkins N.A., Seidman J.G., McMahon A.P., Tabin C.;

RT "Cloning, expression, and chromosomal location of SHH and IHH: two

RL human homologues of the Drosophila segment polarity gene hedgehog.";

RL Genomics 28:44-51(1995).

RN [3]

RP SEQUENCE OF 124-172 FROM N.A.

RX MEDLINE; 95236997.

RA Zhang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,

RA Zhao R., Seldin M.F., Fallon J.F., Beachy P.A.;

RT "Products, genetic linkage and limb patterning activity of a murine

RL hedgehog gene.";

RL Development 120:3339-3353(1994).

CC -1- FUNCTION: INTERCELLULAR SIGNAL. ESSENTIAL FOR A VARIETY OF

CC PATTERNING EVENTS DURING DEVELOPMENT. BINDS TO THE PATCHED (PTC)

CC RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHED MUSCLE (SMO), TO

CC ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IMPLICATED IN

CC ENDOCHONDRAL OSSIFICATION. MAY REGULATE THE BALANCE BETWEEN GROWTH

CC AND OSSIFICATION OF THE DEVELOPING BONES. INDUCES THE EXPRESSION

CC OF PARATHYROID HORMONE-RELATED PROTEIN (PTHrP) (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE

CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE

CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM

CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).

CC -1- TISSUE SPECIFICITY: EXPRESSED IN EMBRYONIC LUNG, AND IN ADULT

CC KIDNEY AND LIVER.

CC -1- PM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY

CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN

CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF

CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-

CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS

CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION

CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE

CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE

CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC -----
 CC EMBL: AB018076; BAA33523.1; -
 CC EMBL: AB018075; BAA33523.1; JOINED.
 CC EMBL: AB010092; BAA33523.1; JOINED.
 CC EMBL: U38517; AAA62178.1; -
 CC MIM: 600726; -
 CC DR PFAM: PF01079; Hint; 1.
 CC DR PFAM: PF01085; HH_Signal; 1.
 CC Developmental protein: Autocatalytic cleavage; Hydrolase; Protease;
 CC Signal.
 CC -----
 CC FT SIGNAL 1 27
 CC FT CHAIN 28 411
 CC FT CHAIN 28 202
 CC FT CHAIN 203 411
 CC FT SITE 202 203
 CC FT SITE 248 248
 CC FT SITE 272 272
 CC FT SITE 272 272
 CC FT ACT_SITE 275 275
 CC FT BINDING 202 202
 CC FT CONFLICT 100 100
 CC FT CONFLICT 244 246
 CC FT CONFLICT 259 259
 CC SQ SEQUENCE 411 AA: 45264 MR: 7888ACB8C0B5647 CRC64;
 CC
 CC Query Match 99.4%; Score 2162; DB 1; Length 411;
 CC Best Local Similarity 99.3%; Pred. No. 3.3e-174;
 CC Matches 408; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 CC
 CC 1 MSPARLRLHFLVLLVLLVPAAMCGGRRVSGRRPRRLVPLAYQGFSPNPEKT 60
 CC 1 MSPARLRLHFLVLLVLLVPAAMCGGRRVSGRRPRRLVPLAYQGFSPNPEKT 60
 CC 1 MSPARLRLHFLVLLVLLVPAAMCGGRRVSGRRPRRLVPLAYQGFSPNPEKT 60
 CC 61 LGASGREKIASNSERFELTPNNDIIFKDEENTGADRLMTQRCRDLNSLAISVMN 120
 CC 61 LGASGREKIASNSERFELTPNNDIIFKDEENTGADRLMTQRCRDLNSLAISVMN 120
 CC 61 LGASGREKIASNSERFELTPNNDIIFKDEENTGADRLMTQRCRDLNSLAISVMN 120
 CC 121 QMVGVLRLVTEGMDDEGHSESLHYEGRAVDITTSRDNRKKGGLARLVEGFDMVY 180
 CC 121 QMVGVLRLVTEGMDDEGHSESLHYEGRAVDITTSRDNRKKGGLARLVEGFDMVY 180
 CC 121 QMVGVLRLVTEGMDDEGHSESLHYEGRAVDITTSRDNRKKGGLARLVEGFDMVY 180
 CC 181 ESKAHVHCYSKSHSAAKTGCGCPAGQVRLSGARVALSAVRPGDRVLAMGDSPTF 240
 CC 181 ESKAHVHCYSKSHSAAKTGCGCPAGQVRLSGARVALSAVRPGDRVLAMGDSPTF 240
 CC 181 ESKAHVHCYSKSHSAAKTGCGCPAGQVRLSGARVALSAVRPGDRVLAMGDSPTF 240
 CC 241 SDVLLFLDREPHRLRAFOVETODPPRLALTPAHLLFTADNHTPEARFATFASHVOP 300
 CC 241 SDVLLFLDREPHRLRAFOVETODPPRLALTPAHLLFTADNHTPEARFATFASHVOP 300
 CC 241 SDVLLFLDREPHRLRAFOVETODPPRLALTPAHLLFTADNHTPEARFATFASHVOP 300
 CC 301 GGVVLAAGVGLDPARAAVSTHVALGAVAPLRKHGTLVVEDVAVASFAAVADHHLAQLA 360
 CC 301 GGVVLAAGVGLDPARAAVSTHVALGAVAPLRKHGTLVVEDVAVASFAAVADHHLAQLA 360
 CC 301 GGVVLAAGVGLDPARAAVSTHVALGAVAPLRKHGTLVVEDVAVASFAAVADHHLAQLA 360
 CC 361 FMPRLRLFSLAMGSGWTPGEGVHWYPLLRLGRLLLESGFHPGLMGSGAGS 411
 CC 361 FMPRLRLFSLAMGSGWTPGEGVHWYPLLRLGRLLLESGFHPGLMGSGAGS 411
 CC 361 FMPRLRLFSLAMGSGWTPGEGVHWYPLLRLGRLLLESGFHPGLMGSGAGS 411
 CC
 CC RESULT 2
 CC IHH_MOUSE
 CC ID IHH_MOUSE STANDARD: PRT: 411 AA.

AC P97812: 061724;
 DT 15-JUL-1999 (rel. 38, Created)
 DT 15-JUL-1999 (rel. 38, Last sequence update)
 DT 15-JUL-1999 (rel. 38, Last annotation update)
 DE INDIAN HEDGEHOG PROTEIN PRECURSOR (IHH) (HHG-2).
 GN IHH.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.
 RC STRAIN-CD-1; TISSUE-KIDNEY;
 RX MEDLINE: 97236802.
 RA Valentini R.P., Brookhiser W.T., Park J., Yang T., Briggs J.,
 RA Dressler G., Holzman L.B.;
 RT "Post-translational processing and renal expression of mouse Indian
 RT hedgehog.";
 RL J. Biol. Chem. 272:8466-8473(1997).
 RN [2]
 RP SEQUENCE OF 76-411 FROM N.A.
 RC STRAIN-C57BL/6J; TISSUE-EMBRYO;
 RX MEDLINE: 94094334.
 RA Echelard Y., Epstein D.J., St Jacques B., Shen L., Mohler J.,
 RA McMahon J.A., McMahon A.P.;
 RT "Sonic hedgehog, a member of a family of putative signaling
 RT molecules, is implicated in the regulation of CNS polarity.";
 RL Cell 75:1417-1430(1993).
 RN [3]
 RP REVISIONS.
 RC STRAIN-C57BL/6J;
 RA St Jacques B.;
 RL Submitted (Apr-1997) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 124-172 FROM N.A.
 RX MEDLINE: 95236997.
 RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,
 RA Zhao R., Seldin M.F., Fallon J.F., Beachy P.A.;
 RT "Products, genetic linkage and limb patterning activity of a murine
 RT hedgehog gene.";
 RL Development 120:3339-3353(1994).
 CC
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT. BINDS TO THE PATCHED (PTC)
 CC RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHED (SMO), TO
 CC ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IMPLICATED IN
 CC ENDOCHONDRAL OSSIFICATION: MAY REGULATE THE BALANCE BETWEEN GROWTH
 CC OF PARATHYROID HORMONE-RELATED PROTEIN (PTHrP).
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: IN THE EMBRYO, DETECTED IN THE DEVELOPING GUT,
 CC AND UROGENITAL SINUS. IN THE ADULT KIDNEY, EPITHELIUM
 CC AND UROGENITAL SINUS. IN THE ADULT KIDNEY, FOUND IN PROXIMAL
 CC CONVOLUTED AND PROXIMAL STRAIGHT TUBULE.
 CC -1- DEVELOPMENTAL STAGE: DETECTED AT 10 DAYS POST COITUM (DPC) IN
 CC DEVELOPING GUT, AT 14.5 DAYS DPC IN THE CARTILAGE PRIMORDIUM AND
 CC IN THE DEVELOPING UROGENITAL SINUS. EXPRESSION INCREASES WITH
 CC GESTATIONAL AGE IN KIDNEY AND DUODENUM, BECOMING MAXIMAL IN
 CC ADULTHOOD.
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

Db 57 LGASGRYEKIAANSERFKELTPNYPNDIIFKDEENTGADRLMTORCKRLNSLAISVNM 116
QY 121 QMPGVKIRVTEGWEDGHSESLHREGRAVDITTSORDRNKYGGLARLAVEGPMVYX 180
Db 117 QMGVVKIRVTEGWEDGHSESLHREGRAVDITTSORDRNKYGGLARLAVEGPMVYX 176
QY 181 ESKAHNCSVKSSESAAKTGCGEPGAGVRLSSGAVNALSAPRGVRLAMEDESPTE 240
Db 177 ESKAHNCSVKSSESAAKTGCGEPGAGVRLSSGAVNALSAPRGVRLAMEDESPTE 236
QY 241 SDVLIETDRPHRLAFVETOTDPPRLATPRAHLFTADNHTPEAPRRATFASHVCP 300
Db 237 SDVLIETDRPHRLAFVETOTDPPRLATPRAHLFTADNHTPEAPRRATFASHVCP 296
QY 301 GQVYLVA-GVPGLOPARVAASVTHVALGATPLTKHGTIVEDVVASCPAAVADHHLAQL 359
Db 297 GHVVLVAVGSGGLQPAEVVGVRRGTGVGAYAPLTHGTLVADVVASCFALVREOQLAQM 356
QY 360 AFMPRLRHFSLANGSWTPRGEGVHWYPOLYRLGLRLLEESFPLGMSGAGS 411
Db 357 AFMPRLRHFSLANGSWTPRGEGVHWYPOLYRLGLRLLEESFPLGMSGAGS 408
RESULT 4
ID IHH_XENLA STANDARD: PRT: 409 AA.
AC 091612;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE INDIAN HEDGEHOG PROTEIN PRECURSOR (IHH) (BAND HEDGEHOG PROTEIN) (X-BHH)
GN IHH OR BHH.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae;
OC Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-EMBRYO;
RX MEDLINE; 95401852.
RA Ekker S.C., McGrew L.L., Lai C.-J., Lee J.J., von Kessler D.P., Moon R.T., Beachy P.A.;
RT "Distinct expression and shared activities of members of the hedgehog gene family of Xenopus laevis";
RL Development 121:2337-2347(1995).
CC -1- FUNCTION: SIGNAL INVOLVED IN THE EARLY INDUCTION AND PATTERNING OF ANTERODORSAL ECTODERM, NERVOUS SYSTEM AND SOMITES. INDUCES ECTOPIC CEMENT GLAND FORMATION IN EMBRYOS.
CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE CELL SURFACE. HEDGEHOG PROTEIN IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC -1- DEVELOPMENTAL STAGE: DETECTABLE WITHIN THE EARLY GASTRULA AT STAGE 14 NEURULA, HIGH EXPRESSION IN NEUROECTODERM, EXPRESSED THROUGHOUT THE NEURAL PLATE AND SUBSEQUENTLY IN BOTH THE NERVOUS SYSTEM AND IN THE DERMATOME OF SOMITES.
CC -1- INDUCTION: BY ACTIVIN.
CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS RESULT IN AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U26404; AAA85165.1;
DR PFAM: PF01079; Hnt; 1.
DR PRINTS: PR00632; SONCHHOG.
KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
KW Signal.
FT SIGNAL 1 23
FT CHAIN 24 409 POTENTIAL.
FT CHAIN 24 197 INDIAN HEDGEHOG PROTEIN.
FT CHAIN 198 409 INDIAN HEDGEHOG PROTEIN N-PRODUCT.
FT SITE 197 198 INDIAN HEDGEHOG PROTEIN C-PRODUCT.
FT SITE 267 267 CLEAVAGE (AUTO-) (BY SIMILARITY).
FT SITE 267 267 INVOLVED IN AUTO-CLEAVAGE (BY SIMILARITY).
FT ACT_SITE 270 270 ESSENTIAL FOR AUTO-CLEAVAGE (BY SIMILARITY).
FT BINDING 197 197 CHOLESTEROL (BY SIMILARITY).
FT DOMAIN 32 35 POLY-ARG.
SQ SEQUENCE 409 AA; 45591 MW; 6FB265367FB98627 CRC64;
Query Match 68.2%; Score 1483; DB 1; Length 409;
Best Local Similarity 70.9%; Pred. No. 3.5e-11;
Matches 285; Conservative 39; Mismatches 74; Indels 4; Gaps 3;
QY 8 PRHFEVLVLLLVYPAAGCGGPRGVGSRPRKLYPLAKOPSPNVEKTLGASGRY 67
Db 4 PXYVLLCAALLLSGAVGCGPGRYVGRRR-PRKLSLSIKQSPNVEKTLGASGRY 62
QY 68 ESKIAANSERFKELTPNYPNDIIFKDEENTGADRLMTORCKRLNSLAISVNMQMPGYKL 127
Db 63 ESKIAANSERFKELTPNYPNDIIFKDEENTGADRLMTORCKRLNSLAISVNMQMPGYKL 122
QY 128 RTEGWEDGHSESLHREGRAVDITTSORDRNKYGGLARLAVEGPMVYXSKAHV 187
Db 123 RTEGWEDGHSESLHREGRAVDITTSORDRNKYGGLARLAVEGPMVYXSKAHV 182
QY 188 CSVKSESAAKTGCGEPGAGVRLSSGAVNALSAPRGVRLAMEDESPTE 247
Db 183 CSVKSESAAKTGCGEPGAGVRLSSGAVNALSAPRGVRLAMEDESPTE 242
QY 248 DREPHRLAFVETOTDPPRLATPRAHLFTADNHTPEAPRRATFASHVCPGYVLA 307
Db 243 DREPHRLAFVETOTDPPRLATPRAHLFTADNHTPEAPRRATFASHVCPGYVLA 302
QY 308 GYPGLOPARVAASVTHVALGATPLTKHGTIVEDVVASCPAAVADHHLAQLAFPLFLF 367
Db 303 NYVGLIPARVSVNQTNGAYAPLTHGTLVADVVASCFALVREOQLAQLVYPLRLV 362
QY 368 HSLANGSWT-PGE--GVHWYPOLYRLGLRLLEESFPLGMSGAGS 406
Db 363 YNDGIITAGTQSQMGHWTSKALYHLGLRLHGNEDFPLGI 404
RESULT 5
ID SHH_MOUSE STANDARD: PRT: 437 AA.
AC 062226;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE SONIC HEDGEHOG PROTEIN PRECURSOR (SHH) (HHG-1).
GN SHH OR HHG1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C57BL/6J; TISSUE-EMBRYO.

RA MEDLINE; 94094334.
RA Echeldard Y., Epstein D.J., St Jacques B., Shen L., Mohler J.,
RA McMahon J.A., McMahon A.P.;
RT "Sonic hedgehog, a member of a family of putative signaling
RT molecules, is implicated in the regulation of CNS polarity.";
RL Cell 75:1417-1430(1993).
RN [2]
RP REVISION TO 122.
RC STRAIN-C57BL/6J;
RA McMahon A.P.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.
RX MEDLINE; 95236997.
RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,
RA Zhao R., Seidman M.F., Fallon J.F., Beachy P.A.;
RT "Products, genetic linkage and limb patterning activity of a murine
RT hedgehog gene.";
RL Development 120:3339-3353(1994).
RN [4]
RP FUNCTION, AND AUTOPROTEOLYTIC CLEAVAGE.
RX MEDLINE; 95254654.
RA Roelink H., Porter J.A., Chiang C., Tanabe Y., Chang D.T.,
RA Beachy P.A., Jessell T.M.;
RT "Floor plate and motor neuron induction by different concentrations of
RT the amino-terminal cleavage product of sonic hedgehog
RT autoproteolysis.";
RL Cell 81:445-455(1995).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 34-195.
RX MEDLINE; 96069744.
RA Hall T.M.T., Porter J.A., Beachy P.A., Leahy D.J.;
RT "A potential catalytic site revealed by the 1.7-A crystal structure of
RT the amino-terminal signaling domain of Sonic hedgehog.";
RL Nature 378:212-216(1995).
RN [6]
RP FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN
RP ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION
CC OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE
CC CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER
CC TARGET, THE GLI ONCOGENE. INTERREGULAR SIGNAL ESSENTIAL FOR A
CC VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED
CC BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE
CC AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE
CC ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH
CC FLOOR PLATE- AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD
CC CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS
CC 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN A NUMBER OF EMBRYONIC TISSUES
CC INCLUDING THE NOTOCHORD, VENTRAL NEURAL TUBE, FLOOR PLATE, LONG
CC BUD, ZONE OF POLARIZING ACTIVITY AND POSTERIOR DISTAL MESENCEPHAL
CC OF LIMBS. IN THE ADULT, EXPRESSED IN LONG AND NEURAL RETINA.
CC -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE DURING GASTRULATION.
CC -1- INDUCTION: BY RETINOIC ACID.
CC -1- PM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
CC C-PRODUCT HAS NO SIGNALING ACTIVITY.
CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
CC -----
CC THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
CC BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION
CC AT THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
CC use BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY

CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL: X76290; CAA53922.1; .
DR	PDB: 1VHH; 2J-UM-96.
DR	MGD; MGI:98287; SHH.
DR	PFAM; PF01079; Hint; 1.
DR	PFAM; PF01085; HH_signal; 1.
DR	PRINTS; PR00632; SONICHOG.
KW	Developmental protein; Autocatalytic cleavage; Hydrolyase; Protease;
KW	Signal; 3D-structure.
FT	SIGNAL
FT	1
FT	24
FT	POTENTIAL.
FT	SONIC HEDEHOG PROTEIN.
FT	SONIC HEDEHOG PROTEIN N-PRODUCT.
FT	SONIC HEDEHOG PROTEIN C-PRODUCT.
FT	CLEAVAGE (AUTO-) (BY SIMILARITY).
FT	INVOLVED IN CHOLESTEROL TRANSFER (BY
FT	SIMILARITY).
FT	INVOLVED IN AUTO-CLEAVAGE (BY
FT	SIMILARITY).
FT	ESSENTIAL FOR AUTO-CLEAVAGE (BY
FT	SIMILARITY).
FT	CHOLESTEROL (BY SIMILARITY).
FT	POLY-GLY.
FT	SEQUENCE
FT	437 AA; 47773 MW; D0EB2F08E7860EF CRC64;
Query Match	58.88; Score 1279.5; DB 1; Length 437;
Best Local Similarity	61.18; Pred. No. 4.6e-100;
Matches	262; Conservative 39; Mismatches 91; Indels 37; Gaps 8;
QY	13 CLVLVL--LLVVPAAWCGPGRVYVSGRRPRPKLVPLAYKQSPNVPEKTLGASGRYEG 69
DB	8 CFIYVLASSLVCPRL-ACGPGRGHG-KRRRPKRLTPLAYQFIPLNVAEKLGLASGRYEG 65
QY	70 KIARSSERFKELTPYNYNDIIFKDEENTGADRLMTQRCCKRNLSTIASVNNQWPGVKLRV 129
DB	66 KITRSERFKELTPYNYNDIIFKDEENTGADRLMTQRCCKRNLSTIASVNNQWPGVKLRV 125
QY	130 TEGWDEDHHSSESHTYGRVADITTSRDNRNKKIGLLARLVEAGFDWVYVESKRNHICS 189
DB	126 TEGWDEDHHSSESHTYGRVADITTSRDNRNKKIGLLARLVEAGFDWVYVESKRNHICS 185
QY	190 VKSEHSAAKAGCGCPAPAOYRLSESGAVALSAPVPRGRLVAMGDSPTFSDVLIFFDR 249
DB	186 VKAESVYAAKSGCGCPGSAIYHLEGGCGTKLYKDLRPGSRVLAAADGQGLYSDLIFFDR 245
QY	250 EPHRLRAEVIETODPPRRLALTPAHLLETA-DNHTEPARFRATFASHVPGQVYLVAG 308
DB	246 DEGAKKVFYVIEETLPPRRLLTAHLLEFVAPHNDSGPTPGSALFASRVPRGQVYVVA 305
QY	309 VFG-----IQPARVAVS-THALGAYALTKHGLIVEDVYASCPAAVAADHLLQDLATWP 363
DB	306 ERGGDRLLPAVHSVTLREEDAGAYAPLTAHGLILIRVYASCAVTEESHWAIRAP 365
QY	364 LRLFSL-----AMGSWPGGVHHYPOLYLRGLLEE 398
DB	366 FRLAALLAALAPARTDGGGGSIPAASATREKGA-EPTAGIHHYSOLLTHIGTWLDS 424
QY	399 GSFHPLGMS 407
DB	425 ETMHPLGMA 433
RESULT	6
SHH_BRARE	STANDARD; PRT: 418 AA.
AC	092008; 013170; 013171;
DT	15-JUL-1999 (Rel. 38, Created)
DT	15-JUL-1999 (Rel. 38, Last sequence update)
DT	15-JUL-1999 (Rel. 38, Last annotation update)
DE	SONIC HEDEHOG PROTEIN PRECURSOR (SHH) (VHH-1).

CC SHH OR VHH1.
 CC Brachydanio rerio (zebrafish) (Zebra danio).
 CC Euxasterota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
 CC Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
 CC Cyprinidae; Cyprinidae; Rasbora; Danio.
 CC [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE-EMBRIO.
 RC MEDLINE: 94170375.
 RA Roelink H., Augsburger A., Heemskerk J., Korzh V., Norlin S.,
 RA Ruiz I Altaba A., Tanabe Y., Placzek M., Edlund T., Jessell T.M.,
 RA Dodd J.,
 RT "Floor plate and motor neuron induction by vhh-1, a vertebrate
 RT homolog of hedgehog expressed by the notochord.";
 RL Cell 76:761-775(1994).
 RN [2]
 RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.
 RX MEDLINE: 96014264.
 RA Eker S.C., Unger A.R., Greenstein P., von Kessler D.P., Porter J.A.,
 RA Moon R.T., Beachy P.A.;
 RT "Patterning activities of vertebrate hedgehog proteins in the
 RT developing eye and brain.";
 RL Curr. Biol. 5:944-955(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 96083328.
 RA Fietz M.J., Concordet J.-P., Barbosa R., Johnson R., Krauss S.,
 RA McMahon A.P., Tabin C., Ingham P.W.;
 RT "The hedgehog gene family in Drosophila and vertebrate development.";
 RL Development Suppl. 43-51(1994).
 RN [4]
 RP SEQUENCE OF 30-92 AND 113-170 FROM N.A.
 RC TISSUE-MUSCLE;
 RX MEDLINE: 97075114.
 RA Zardoya R., Abounef E., Meyer A.;
 RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species
 RT closely related to the zebrafish.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED BY THE
 CC NOTOCHORD THAT INDUCES SOMITE PATTERNING, DORSO-VENTRAL PATTERNING
 CC OF THE BRAIN AND EARLY PATTERNING OF THE DEVELOPING EYES. DISPLAYS
 CC RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO
 CC ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH,
 CC PTC REPRSESSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE VENTRAL MIDLINE OF THE NEURAL
 CC TUBE AND BRAIN. ALSO FOUND IN THE NOTOCHORD AND IN DEVELOPING FIN
 CC BUD. IN THE DEVELOPING BRAIN, EXPRESSION OCCURS IN DOMAINS THAT
 CC INCLUDE A DISCRETE REGION IN THE FLOOR OF THE Diencephalon
 CC -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE IN THE INNER CELL LAYER OF
 CC THE EMBRYONIC SHIELD DURING GASTRULATION. BY 9.5 HRS OF
 CC DEVELOPMENT, EXPRESSED IN A CONTINUOUS BAND THAT EXTENDS FROM THE
 CC TAIL TO THE HEAD. THE ANTERIOR BOUNDARY OF EXPRESSION BEING
 CC POSITIONED IN THE CENTER OF THE ANIMAL POLE ANTERIOR TO THE
 CC PRESUMPTIVE MIDBRAIN.
 CC -1- PTH: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L27585; AAA20998.1; -
 CC DR EMBL: U30711; AAC59742.1; -
 CC DR EMBL: 235669; CA84738.1; -
 CC DR EMBL: U51351; AAB38575.1; -
 CC DR EMBL: U51370; AAB38593.1; -
 CC DR ZFIN: ZDB-GENE-980526-166; SHH.
 CC DR PFM: PF01079; Hint: 1.
 CC DR PFM: PF01085; HH_signal; 1.
 CC DR PRINTS: PR00632; SONICHHOG.
 CC KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 CC Signal.
 CC FT SIGNAL 1 23
 CC FT CHAIN 24 418
 CC FT CHAIN 24 197
 CC FT CHAIN 198 418
 CC FT SITE 197 198
 CC FT SITE 243 243
 CC FT SITE 267 267
 CC FT ACT_SITE 270 270
 CC FT BINDING 197 197
 CC FT SEQUENCE 418 AA; 46402 MW; CF000A9FFD2F5795 CRC64;
 CC
 CC Query Match 58.6%; Score 1274; DB 1; Length 418;
 CC Best Local Similarity 61.2%; Pred. No. 1.3e-99;
 CC Matches 254; Conservative 46; Mismatches 101; Indels 14; Gaps 4;
 CC
 CC QY 5 RLRLPRLHCLVLLLVVPAAMGCGPGRVSGRRPRKLVPLAYKQSPNPKYTGAS 64
 CC DB 2 RLRLVRLVLSLTSLVY-SGLACGPGRGY-RRRHKKLPLPKYKQFIPVAKRTYGAS 59
 CC QY 65 GRYGKTLRSERKELTPNPNPDIKDEENTGADSLMORCDRLNSLAISYMNQWP 124
 CC DB 60 GRYGKTLRSERKELTPNPNPDIKDEENTGADSLMORCDRLNSLAISYMNQWP 119
 CC QY 125 VKLVTGMDGDGHHSESLHYEGRAVDITSDRDKNKYGILLARLAYEAGDQWYYESKA 184
 CC DB 120 VKLVTGMDGDGHHSESLHYEGRAVDITSDRDKNKYGILLARLAYEAGDQWYYESKA 179
 CC QY 185 HVHGSVKSSESAAKKTGGCPFAGQVRLSAGARVALSAVRGDRVLAMGEDGSPFESDVL 244
 CC DB 180 HHSVKAKENSVAKSGCGFPGSALVSLDGGQKAVVDLNDGKVLAAADSGNLFVSDFI 239
 CC QY 245 IFLRPRERLAPQVITPDPRRLATLPAHLFTADNHPARFATFASHQPOQVY 304
 CC DB 240 MFTDSDSTRVVFVIEQPEVETITLPAHLFLVDNSTDHLMTAAASVSRQKV 299
 CC QY 305 LVAGVPG-LQPARVAASVTHVALGAPLTKGLTVEDVVAASCFAAVADHHLQLFWP 363
 CC DB 300 MVDVDSGLKSVIQRITTEORGSFAPVYHGIIVDRIILASCIYAIEDOGLAHLAFAP 359
 CC QY 364 LRLH-----SLAWGWTGEGVHWYPOLLYRLGLHLEGGSPFLGNS 407
 CC DB 360 ARLYVYSSFLPONSRSRNATLQGGVHWYSRLYQMGWGLDLSNMLHPLGNS 414
 CC
 CC RESULT 7
 CC SHH_CHICK STANDARD; PRT; 425 AA.
 CC AC 091035;
 CC DT 15-JUL-1999 (Rel. 38, Created)
 CC DT 15-JUL-1999 (Rel. 38, Last sequence update)
 CC DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE SONIC HEDGEHOG PROTEIN PRECURSOR (SHH).
GN SHH.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
OC Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LIMB BUD.
RX MEDLINE: 94094333.
RA Riddle R.D., Johnson L., Laufer E., Tabin C.;
RT "sonic hedgehog mediates the polarizing activity of the zpa."; Cell 75:1401-1416(1993).
RL [2]
RN FUNCTION AND AUTOPROTEOLYTIC CLEAVAGE.
RX MEDLINE: 95254654.
RA Roelink H., Porter J.A., Chiang C., Tanabe Y., Chang D.T.,
RA Beachy P.A., Jessell T.M.;
RT "Floor plate and motor neuron induction by different concentrations of the amino-terminal cleavage product of sonic hedgehog autoproteolysis."; Cell 81:445-455(1995).
RL [-1]
CC FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHED (SMO), TO ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH FLOOR PLATE AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION.
CC [-1]
CC SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC [-1]
CC TISSUE SPECIFICITY: EXPRESSED IN THE POSTERIOR LIMB BUD MESENCHYME, THE HENSEN'S NODE, THE NOTOCHORD, AND THE FLOOR PLATE OF THE NEURAL TUBE.
CC [-1]
CC DEVELOPMENTAL STAGE: FIRST DETECTABLE AT STAGE 17 DURING THE INITIATION OF LIMB BUD FORMATION. FROM THAT POINT ONWARDS, THE EXPRESSION PATTERN EXACTLY MATCHES THE LOCATION OF THE ZONE OF POLARIZING ACTIVITY (ZPA).
CC [-1]
CC INDUCTION: BY RETINOID ACID.
CC [-1]
CC PTC: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE C-PRODUCT HAS NO SIGNALING ACTIVITY.
CC [-1]
CC SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announce/](http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch) or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L28099; AAA72428.1; -
DR PRAM; PF01079; Hint; 1.
DR PRAM; PF01085; HH_signal; 1.
DR PRINTS; PR00632; SONICHHOG.
KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
KM Signal.
FT SIGNAL. 1 26 POTENTIAL.
FT CHAIN 27 425 SONIC HEDGEHOG PROTEIN.
FT CHAIN 27 200 SONIC HEDGEHOG PROTEIN N-PRODUCT.

FT CHAIN 201 425
FT SITE 200 201
FT SITE 246 246
FT SITE 270 270
FT ACT_SITE 273 273
FT BINDING 200 200
FT DOMAIN 390 393
SQ SEQUENCE 425 AA; 46474 MW; DA627443DA40173 CRC64;
Query Match 58.4%; Score 1270.5; DB 1; Length 425;
Best Local Similarity 63.0%; Pred. No. 2,6e-99;
Matches 257; Conservative 38; Mismatches 96; Indels 17; Gaps 5;
QY 20 LVPAANGCGFRVYSGRRPRRLVPLAKQSPNYPETLGSAGRYEGKIANSSEFEK 79
DB 19 LVSSGLTCGPGRGIG-KRRHPKRLPLAYKQFIPNVAERTIGASGRYEGRITNSSEFEK 77
QY 80 ELTPNYPDITFKDEMTGADRLMTORCKDLNLSATSVNMOWGVKLRVTEGDEGHH 139
DB 78 ELTPNYPDITFKDEMTGADRLMTORCKDLNLSATSVNMOWGVKLRVTEGDEGHH 137
QY 140 SEESLHYEGRADVITTSDRNRKYGILLRLAVEAGFDMVYESKARHCVSEKSAAK 199
DB 138 SEESLHYEGRADVITTSDRNRKYGILLRLAVEAGFDMVYESKARHCVSEKSAAK 197
QY 200 TGGCFPAGAOVRLSGARVALSAVRPQDVLAMGEDSPFTSDVLTILDRPHRLAPQV 259
DB 198 SGGCFPSATVHEHGRTKLVKDLSPGDRVLAADADRLYSDLTFLDRMDSRKLFYV 257
QY 260 IETQDPRLALFAHLFTADNTEBAA----REFATPSHVPGQYVAVAGPGQ-- 314
DB 258 IETQDPRLALFAHLFTADNTEBAA----REFATPSHVPGQYVAVAGPGQ-- 317
QY 314 PARVAAS-THVALGAYAPLTKHGTLVEDVASCFAVADVADHILQALFPLRLFHSIAW 372
DB 318 PASVHSVLSREASGAYAPLTAQGTILLNRVLAACVAVIEHSHANAFAPRLAQGLLA 377
QY 373 G-----SWTPEGVHTVPLQYRLGLLEGGFHPGSGAGS 411
DB 378 ALCPDGAIPRTATTTGTGIIHWYSRLYRIGSVWLDGDLHPLGMVAPAS 425
RESULT 8
SHH_HUMAN STANDARD; PRT; 462 AA.
AC Q15465;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE SONIC HEDGEHOG PROTEIN PRECURSOR (SHH) (HHG-1).
GN SHH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-FETAL LUNG;
RX MEDLINE: 96070431.
RA Marigo V., Roberts D.J., Lee S.M.K., Tsukurov O., Levi T.,
RA Gaster J.M., Epstein D.J., Gilbert D.J., Copeland N.G., Seidman C.E.,
RA Jenkins N.A., Seidman J.G., McMahon A.P., Tabin C.;
RT "Cloning, expression, and chromosomal location of SHH and HH: two human homologues of the Drosophila segment polarity gene hedgehog."; Genomics 28:44-51(1995).
RL [2]
RN SEQUENCE OF 1-187 FROM N.A.
RA Strong C., Graves T., Suterer C., Ozersky P.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [3]

RP SEQUENCE OF 119-167 FROM N.A.
 RX MEDLINE: 95236997.
 RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Sliamandl B.K.,
 Zhao R., Seldin M.F., Fallon J.F., Beachy P.A.;
 RT "Products, genetic linkage and limb patterning activity of a murine
 RL hedgehog gene";
 RN Development 120:3339-3353(1994).
 (4)
 RP VARIANTS HPE3 ARG-31; GLY-117 AND ARG-117.
 RX MEDLINE: 97051937.
 RA Roesler E., Belloni E., Gaudenz K., Jay P., Berta P., Scherer S.W.,
 Tsui L.-C., Muenke M.;
 RT "Mutations in the human Sonic Hedgehog gene cause holoprosencephaly";
 RL Nat. Genet. 14:357-360(1996).
 CC -1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN
 CC ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION
 CC OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESENTS THE
 CC CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER
 CC TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A
 CC VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED
 CC BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE
 CC AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE
 CC ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH
 CC FLOOR PLATE- AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD
 CC CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS
 CC 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN FETAL INTESTINE, LIVER, LUNG, AND
 CC KIDNEY. NOT EXPRESSED IN ADULT TISSUES.
 CC -1- PTV: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- DISEASE: DEFECTS IN SHH ARE THE CAUSE OF THE AUTOSOMAL DOMINANT
 CC DISORDER HOLOPROSENCEPHALY TYPE 3 (HPE3). HPE3 IS A GENETICALLY
 CC HETEROGENEOUS DISEASE THAT AFFECTS THE MIDLINE DEVELOPMENT OF THE
 CC FOREBRAIN AND MIDFACE. HPE IS ASSOCIATED WITH SEVERAL DISTINCT
 CC FACIES AND PHENOTYPIC VARIABILITY. IN THE MOST EXTREME CASES,
 CC ANOPHTHALMIA OR CYCLOPIA IS EVIDENT ALONG WITH A CONGENITAL
 CC ABSENCE OF THE MATURE NOSE. THE LESS SEVERE FORM FEATURES FACIAL
 CC DYSMORPHIA CHARACTERIZED BY OCULAR HYPERTELORISM, DEFECTS OF THE
 CC UPPER LIP AND/OR NOSE, AND ABSENCE OF THE OLFACTORY NERVES OR
 CC CORPUS CALLOSUM.
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L38518; AA662179.1; -
 CC EMBL: AC002484; AAB67604.1; -
 CC MIM: 600725; -
 CC MIM: 142945; -
 CC PFAM: PF01079; Hint: 1.
 CC PFAM: PF01083; HH_signal; 1.
 CC PRINTS: PR00632; SONICHHOG.
 CC Developmental protein: Autocatalytic cleavage; Hydrolase; Protease;
 CC Signal: Disease mutation.
 CC SIGNAL 1 23 POTENTIAL.
 CC CHAIN 24 462 SONIC HEDGEHOG PROTEIN.

FT CHAIN 24 197 SONIC HEDGEHOG PROTEIN N-PRODUCT.
 FT CHAIN 198 462 SONIC HEDGEHOG PROTEIN C-PRODUCT.
 FT SITE 197 198 CLEAVAGE (AUTO-) (BY SIMILARITY).
 FT SITE 243 243 INVOLVED IN CHOLESTEROL TRANSFER (BY
 FT SITE 267 267 SIMILARITY).
 FT SITE 267 267 INVOLVED IN AUTO-CLEAVAGE (BY
 FT ACT_SITE 270 270 SIMILARITY).
 FT BINDING 197 197 ESSENTIAL FOR AUTO-CLEAVAGE (BY
 FT DOMAIN 407 411 CHOLESTEROL (BY SIMILARITY).
 FT VARIANT 31 31 POLY-GLY.
 FT VARIANT 117 117 G -> R (IN HPE3).
 FT VARIANT 117 117 /FTID-VAR_003619.
 FT VARIANT 117 117 W -> G (IN HPE3).
 FT VARIANT 117 117 /FTID-VAR_003620.
 FT VARIANT 117 117 W -> R (IN HPE3).
 FT VARIANT 117 117 /FTID-VAR_003621.
 FT SEQUENCE 462 AA: 49607 MM: DD687AF582A4749 CRC64;
 SQ
 Query Match 58.3%; Score 1268.5; DB 1; Length 462;
 Best Local Similarity 58.3%; Pred. No. 4.2e-99;
 Matches 264; Conservative 45; Mismatches 85; Indels 59; Gaps 9;
 QY 13 CLVLLLL--LYVPAAMCGGPGRVGSSRRPRKLVPLAYKQFSRVPEKTLGASGRYEGK 70
 DB 7 CLLVLVSSLVLCGSLAGPGRGFG-KRRHPRKLTPLVYKQFIPVAKETLGASGRYEGK 65
 QY 71 IARSSERKELTPNNPPIIFKDEENTGADRLMTORCDRLNSLAISYNNOMPVKLYRT 130
 DB 66 ISRNERKELTPNNPPIIFKDEENTGADRLMTORCDRLNSLAISYNNOMPVKLYRT 125
 QY 131 EGWDEGHHSESLHYEGRAVDITTSDBRNNKYGILARLAVEAGEFVYYSKAAHVCSV 190
 DB 126 EGWDEGHHSESLHYEGRAVDITTSDBRNNKYGILARLAVEAGEFVYYSKAAHVCSV 185
 QY 191 KSEHNAAKTGCGCPAGQVQLSARVALSAVRPGDVLAMGEGSTFSDVLIFFLDRE 250
 DB 186 KAENSVAAKSGCGFSGATVHLDEGGTKLVKDLISGDRVLAADOGRLYSDFLFLDRD 245
 QY 251 PHRLAFVFIETDPDRRLATPLPHLFTFA--DWHF-EPAARF-----RATF 294
 DB 246 DGAKVFFVETREKRELLTAHLFLVAAPHNSATCEPFASSGSGPGSCALGRALF 305
 QY 295 ASHVOPGGYLVV---GVPGIOPARVAAS-THVALGAVAPLTKHGLTVVEDVVASCF 349
 DB 306 ASRVPRGGRYVVAERDGRLLPLAAVHSTVLSSEAAGAVAPLTAQGTLINRYLASCY 365
 QY 350 AYADHHLAQLFWPLRLFHSL-----AMGSWTPG 378
 DB 366 VIEHSMVHRAFAFPRRLAHLAALAPARTDRGDSGGGGRGGRVALTPGAADAPG 425
 QY 379 ---EGVHVPOLLYRLGLRLLLEGSFHLGMS 407
 DB 426 AGATAGITMYSQLYQIGTWTLLDSEALHPLGMA 458
 RESULT 9
 SHH_CYPNY
 ID SHH_CYPNY STANDARD; PRT; 432 AA.
 AC Q90385;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE SONIC HEDGEHOG PROTEIN PRECURSOR (SHH).
 GN SHH.
 OS Cynops pyrrhogaster (Japanese common newt).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Amphibia;
 CC Batrachia; Caudata; Salamandridae; Cynops.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-EMBRIO;
 RX MEDLINE: 96136334.

RA Takahashi T., Takahashi T.C., Inoue K., Ogawa M., Takeshima K.:
 RT "Activation of two Cypops genes, fork head and sonic hedgehog, in
 RL animal cap explants.";
 CC Blochem. Biophys. Res. Commun. 218:395-401(1996).
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT AND METAMORPHOSIS. INVOLVED
 CC IN LIMB FORMATION, PATTERNING OF THE CENTRAL NERVOUS SYSTEM AND
 CC VENTRAL SOMITE DIFFERENTIATION. INDUCES ECTOPIC CEMENT GLAND
 CC FORMATION IN EMBRYOS. BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH
 CC FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE
 CC TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH, PTC
 CC REPRESSSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- INDUCTION: ACTIVATED BY ACTIVIN, BASIC FIBROBLAST GROWTH FACTOR
 CC (bFGF) AND FGF.
 CC -1- PFM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; D63339; BAA09657.1; -;
 DR PFM; PFO1079; HInt; 1.
 DR PFM; PFO1085; HR.Signal; 1.
 DR PRINTS; PRO0632; SONICHOG.
 KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 KW Signal.
 FT SIGNAL. 1 26
 FT CHAIN 27 432
 FT CHAIN 27 200
 FT CHAIN 201 432
 FT SITE 200 201
 FT SITE 268 268
 FT SITE 268 268
 FT ACT_SITE 271 271
 FT BINDING 200 200
 FT SEQUENCE 432 AA; 47847 MW; B453C7E746C8E5A8 CRC64;
 SO
 Query Match 58.1%; Score 1263; DB 1; Length 432;
 Best Local Similarity 60.4%; Pred. No. 1.le-98;
 Matches 252; Conservative 48; Mismatches 89; Indels 28; Gaps 6;
 QY 20 LVPAMCGGGRVVGSRPRPKLPLVAYKQSPNVEKTLGASRGYKJARSSEK 79
 DB 19 LTVSGSGSCGGRGIGTKRR-FKRLPLAYKQFTPNVEKTLGASRGYKJARSSEK 77
 QY 80 ELTPYNDIIFKDEENGGARLMTORCKDRNLNSLAISVMNMPVKLRVTEGMEDGH 139
 DB 78 ELTPYNDIIFKDEENGGARLMTORCKDRNLNSLAISVMNMPVKLRVTEGMEDGH 137
 QY 140 SEESLHYGRAVDITTSRDRNRKYLGLARLVAEAGFDWYYSKKAHVCSSESAAK 199
 DB 138 FEESLHYGRAVDITTSRDRNRKYLGLARLVAEAGFDWYYSKKAHVCSSESAAK 197

QY 200 TGCGPAGAGVRLSESGARVNLASVPGDRVLAMEGDSPTRESVDLITLDEPHRLRAFOV 259
 DB 198 SGGCPGASATVALLAQGRITPKDLRPGDRVLADGLGLVSDLEDFMDKREYRKVFYV 257
 QY 260 IETODPPRLRLTAHLLETFADN-----TEPAARFATFASHVPGQYVLVAG 308
 DB 258 IETS--KERVLTAAHLFLVQALPBGDSGDFSVSGSAGFRMFASVAGRVLTV 315
 QY 309 VP--GLQPARVAASVTHALGAVAPLTKHGLVVEDVAVSCFAAVADHHLAQLAEFWPLRL 366
 DB 316 REGGLREATVERVLEBATGAVPVTAHGVIVDRVLASCYAVIEHSHMAWAFAPLRY 375
 QY 367 -----FHSIANGSWP-----GEVHYRQQLYRGLLLEGSFPLMSGAGS 411
 DB 376 GLGITLFFSPDYSHPDYSHPAPDSQSEGVHWSSEILYRIGTWTLODTHPIGMAKSS 432
 RESULT 10
 SHH_XENLA STANDARD; PRT; 444 AA.
 ID SHH_XENLA
 AC Q92000; Q91894;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-FEB-2000 (Rel. 39, Last annotation update)
 DE SONIC HEDGEHOG PROTEIN PRECURSOR (X-SHH) (VHH-1).
 GN SHH.
 OS Xenopus laevis (African clawed frog).
 OC Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
 OC Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae; Xenopodinae;
 OC Xenopus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-INTESINE;
 RX MEDLINE; 95357169.
 RA Stolor M.A., Shi Y.-B.;
 RT "Xenopus sonic hedgehog as a potential morphogen during embryogenesis
 RT and thyroid hormone-dependent metamorphosis.";
 RL Nucleic Acids Res. 23:2555-2562(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-EMBRYO;
 RX MEDLINE; 95401852.
 RA Ekker S.C., McGrew L.L., Lai C.-J., Lee J.-J., von Kessler D.P.,
 RA Moon R.T., Beachy P.A.;
 RT "Distinct expression and shared activities of members of the hedgehog
 RT gene family of Xenopus laevis.";
 RL Development 121:2337-2347(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE-NOTOCHORD;
 RX MEDLINE; 96028338.
 RA Ruiz i Altaba A., Jessell T.M., Roelink H.;
 RT "Restrictions to floor plate induction by hedgehog and winged-helix
 RT genes in the neural tube of frog embryos.";
 RL Mol. Cell. Neurosci. 6:106-121(1995).
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT AND METAMORPHOSIS. INVOLVED
 CC IN LIMB FORMATION, PATTERNING OF THE CENTRAL NERVOUS SYSTEM AND
 CC VENTRAL SOMITE DIFFERENTIATION. INDUCES ECTOPIC CEMENT GLAND
 CC FORMATION IN EMBRYOS. BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH
 CC FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE
 CC TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH, PTC
 CC REPRESSSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN NOTOCHORD AND NEURAL
 CC FLOOR PLATE DURING EMBRYOGENESIS. IN TADPOLE, HIGH EXPRESSION IS
 CC OBSERVED IN PANCREAS/STOMACH, MODERATE EXPRESSION IN TAIL, AND LOW
 CC EXPRESSION IN INTESTINE, BRAIN, AND HIND LIMB.
 CC -1- DEVELOPMENTAL STAGE: FIRST DETECTED AT THE NEURULA (STAGES 16-17).


```

Db 63 GKTEGKTRRSEKELTPYNDIIFKDEENTGADRLMKTCKDKNSLAISVMNMPG 122
125 VKLRVTEGMDDEGHSESLHYEGRAVDITTSRDRNKYGLARLAVEAGFDWVYESKA 184
123 VKLRVTEGMDDEGHSESLHYEGRAVDITTSRDRNKYGLARLAVEAGFDWVYESKA 182
185 HVYCSKSEHSAAKTGCGCPAGAVRLESGARVALSAPGGRVLAAMEGDSPTSDVL 244
183 HHCYSKAEHSAKSGCGCPAGAVRLESGARVALSAPGGRVLAAMEGDSPTSDVL 242
245 IFDREPHRLRAFOVETODPPRRLATPLAHLFTADNHTEPARFAPRATFASHVOGYV 304
243 MFDHDPTRKROFVIETSEPFKLTITAAHLVF--GNSAASGATITATSNVKEFDIV 300
305 LV--AGVPGLOPARVAVSTHVALGAVAPLTKGTLVEDVAVASCPAAVDHHLAQLAFW 362
301 LVWEDTCESLSTYKRIYEEHSGSPAPVTAGTIVDQVLASCYAVIENHKMAHMAFA 360
363 PLALFSLANGSWT-----PGEVHWYPOLYRLGRLLLEESGFHPLGM 406
361 PVRLCRKLK--TWLFPARESNVNFQEDGIHWSNMLFHIGSWLLDRDSFHLGI 412

RESULT 13
ID IHH_BRARE STANDARD: PRT: 412 AA.
AC Q98862;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE INDIAN HEDGEHOG PROTEIN PRECURSOR (IHH) (ECHIDNA HEDGEHOG PROTEIN).
GN IHH OR EHH.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
OC Cyprinoidae; Cyprinidae; Rasbora; Danio.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 96310864.
RA Currie P.D., Ingham P.W.;
RT "Induction of a specific muscle cell type by a hedgehog-like protein
in zebrafish."
RL Nature 382:452-455(1996).
RN [2]
RP SEQUENCE OF 113-170 FROM N.A.
RX TISSUE-MUSCLE:
RA Zardoya R., Abouheif E., Meyer A.;
RT "Evolutionary analyses of hedgehog and hoxd-10 genes in fish species
closely related to the zebrafish."
RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).
CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
PATTERNING EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED BY THE
NOTOCHORD THAT INDUCES SOMITE PATTERNING AND MUSCLE PIONEER
DIFFERENTIATION.
CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN THE NOTOCHORD.
CC -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE AT THE MID-GASTRULA STAGE.
DISAPPEARS AT THE END OF THE SOMITOGENESIS.
CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb.ch).
CC
DR EMBL: Y08426; CAA69702.1; -.
DR ZFIN: ZDB-GENE-980526-135; EHH.
DR PIR: P01079; Hint: 1.
DR PIR: P01085; HH_Signal: 1.
DR PRINTS: PR00632; SONICHHOG.
KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
KW Signal.
FT SIGNAL 1 23
FT CHAIN 24 412
FT CHAIN 24 197
FT CHAIN 198 412
FT SITE 197 198
FT SITE 197 198
FT SITE 246 246
FT SITE 269 269
FT SITE 272 272
FT ACT_SITE 272 272
FT BINDING 197 197
FT CONFLICT 121 121
FT SEQUENCE 412 AA: 45746 MW: 2104350528278CAL CRC64:

Query Match 55.0%; Score 1196; DB 1; Length 412;
Best Local Similarity 61.7%; Pred. No. 4.4e-93;
Matches 258; Conservative 35; Mismatches 93; Indels 32; Gaps 13;

9 RHFCVLVL---LLVVPAMGCGPGRVGSRRPRKRVPLATYKQSPNVYEKTLGASG 65
Db 2 RUSTAAALTLGFLAFSPAYDGGPGRGYG-KRRPRKRLPLATYKQSPNVYEKTLGASG 60
QY 66 RYEGKIRARSSERKELTPYNDIIFKDEENTGADRLMKTCKDKNSLAISVMNMPGV 125
Db 61 RYEGKIRARSSERKELTPYNDIIFKDEENTGADRLMKTCKDKNSLAISVMNMPGV 120
QY 126 KLRVTEGMDDEGHSESLHYEGRAVDITTSRDRNKYGLARLAVEAGFDWVYESKAH 185
Db 121 KLRVTEGMDDEGHSESLHYEGRAVDITTSRDRNKYGLARLAVEAGFDWVYESKAH 180
QY 186 HVYCSKSEHSAAKTGCGCPAGAVRLESGARVALSAPGGRVLAAMEGDSPTSDVL 244
Db 181 HVYCSKSEHSAAKTGCGCPAGAVRLESGARVALSAPGGRVLAAMEGDSPTSDVL 240
QY 243 VLFIDREPHRLRAFOVETODPPRRLATPLAHLFTADNHTEPARFAPRATFASH 297
Db 241 VLFIDREPHRLRAFOVETODPPRRLATPLAHLFTADNHTEPARFAPRATFASH 299
QY 298 VOPGQVLYAVGLOPARVAVSTHVAL---GAVAPLTKGTLVEDVAVASCPAAVD 353
Db 300 AVGQCCILL---GKRRKFSOI-THVGREDQGLPPPLTAGTIVVNDVLSYCAAVNR 354
QY 354 HHQAQAFPLFLFSLANGSWT-PGE---GVHWYPOLYRLGRLLLEESGFHPLGM 406
Db 355 QRLAHMAFAPRLRLY---SWTGPQVLKNGLIHWSOVLIGLKLDSLELHPLAL 407

RESULT 14
ID DHH_MOUSE STANDARD: PRT: 396 AA.
AC Q61488;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE DESERT HEDGEHOG PROTEIN PRECURSOR (DHH) (HHS-3).

```


GN DHH.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/5V.
 RX MEDLINE: 94094334.
 RA Echelard Y., Epstein D.J., St Jacques B., Shen L., Mohler J.,
 RA McMahon J.A., McMahon A.P.;
 RT "Sonic hedgehog, a member of a family of putative signaling
 RT molecules, is implicated in the regulation of CNS polarity.";
 RL Cell 75:1417-1430(1993).
 RN [2]
 RN SEQUENCE OF 120-168 FROM N.A.
 RX MEDLINE: 95236997.
 RA Zhang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,
 RA Zhao R., Seldin M.F., Fallon J.F., Beachy P.A.;
 RT "Products, genetic linkage and limb patterning activity of a murine
 RT hedgehog gene.";
 RL Development 120:3339-3353(1994).
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT. MAY FUNCTION AS A
 CC SPERMATOCYTE SURVIVAL FACTOR IN THE TESTES. ESSENTIAL FOR TESTES
 CC DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ADULT TESTES. NOT EXPRESSED IN
 CC LAMB BUDS.
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING. WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X76292; CAA53924.1; -
 DR MGI: 94891; DHH.
 DR PFAM: PF01079; HInt: 1.
 DR PFAM: PF01085; HH.Signal: 1.
 DR PRINTS: PRO0632; SONICHHOG.
 KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 KW Signal.
 FT SIGNAL: 1 22
 FT CHAIN: 23 396
 FT CHAIN: 23 198
 FT CHAIN: 199 396
 FT SITE: 198 199
 FT SITE: 244 244
 FT SITE: 268 268
 FT ACT SITE: 271 271
 FT BINDING: 198 198
 FT SEQUENCE: 396 AA; 43542 MW; AFE6051BE950FD8 CRC64;
 Query Match 54.0%; Score 1175; DB 1; Length 396;

Best Local Similarity 58.5%; Pred. No. 2.4e-91;
 Matches 234; Conservative 50; Mismatches 102; Indels 14; Gaps 5;
 QY 3 PARLRRLHFCVLLVLLVPAAMGCGPR-VYSGRRPRPRKPLVAYQFSPNVKTL 61
 DB 4 PASLPLC--CLALLA-----SAQSGPRGPGVRRYRKQVLLYQFSPNERTL 57
 QY 62 GASRGYKIANSSERFELTPNYNDIIFKDEENTGADRLMTQRCCKRLNSLAISVNO 121
 DB 58 GASGAEAGHTVSGSERFRLVYNYNDIIFKDEENSGADRLMTQRCCKRLNSLAISVNM 117
 QY 122 WPGVRLVTEGMDHDESHSESLHYEGRAVDITTSRDKNKGLARLAVEGFMVYVE 181
 DB 118 WPGVRLVTEGMDHDESHSESLHYEGRAVDITTSRDKNKGLARLAVEGFMVYVE 177
 QY 182 SKAHVCSYKSEHSAAKGCGEPAGAGVRLSEGAVALSAVPCGVYLMKEDSGSPTS 241
 DB 178 SRNHIVSVKADNSLAVRAGCGEPGATVRLASGERKGLRELHGRGWVLAADAAGVPT 237
 QY 242 DVILFIDREPHRLARFQVLETQDPPRLALTPAHLFTADNHTPEARFRATFASHVPG 301
 DB 238 PVLLFDRLQRRASFVAVETERPRKLLTPMHLVFAARGPAPAGDPAPVFAFRLRAG 297
 QY 302 QYLVAGVGLQPARVAASTHVALGAVAPLKHGTLVYEDVVASCFEAVADHHLAQLAF 361
 DB 298 DSVLAPGADLPARVAVRAREAVGFAPLTAHGTLLVNDVASCYAVLESHQWHRAF 357
 QY 362 WPLRLFHSLSAMGSMTPG-----EGVHWYPLLRYRLGLLL 396
 DB 358 APLRLHAL--GALDPGAVPGMHWYRLRYRLAEELM 395
 RESULT 15
 ID DHH HUMAN STANDARD: PRT: 396 AA.
 AC 043323; Q15794;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE DESERT HEDGEHOG PROTEIN PRECURSOR (DHH) (HNG-3).
 GN DHH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tate G., Endo Y., Mitsuura T.;
 RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RN SEQUENCE OF 85-178 FROM N.A.
 RC TISSUE=KIDNEY;
 RA Drummond I.A.;
 RT "Human desert hedgehog.";
 RT Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT. MAY FUNCTION AS A
 CC SPERMATOCYTE SURVIVAL FACTOR IN THE TESTES. ESSENTIAL FOR TESTES
 CC DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING. WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AB010994; BAA24866.1; -
DR EMBL; AB010581; BAA24866.1; JOINED.
DR EMBL; AB010993; BAA24866.1; JOINED.
DR EMBL; U59748; AAB03398.1; -
DR PRAM; PF01079; Hntc; 1.
DR PRAM; PF01085; HH_signal; 1.
DR PRINTS; PR00632; SONIC_HOG.
KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
KW Signal.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 396 DESERT HEDGEHOG PROTEIN.
FT CHAIN 23 198 DESERT HEDGEHOG PROTEIN N-PRODUCT.
FT CHAIN 199 396 DESERT HEDGEHOG PROTEIN C-PRODUCT.
FT SITE 198 199 CLEAVAGE (AUTO-) (BY SIMILARITY).
FT SITE 244 244 INVOLVED IN CHOLESTEROL TRANSFER (BY
FT SITE 268 268 SIMILARITY).
FT SITE 268 268 INVOLVED IN AUTO-CLEAVAGE (BY
FT ACT_SITE 271 271 SIMILARITY).
FT BINDING 198 198 ESSENTIAL FOR AUTO-CLEAVAGE (BY
FT CONFLICT 177 177 SIMILARITY).
FT CONFLICT 177 177 CHOLESTEROL (BY SIMILARITY).
SQ SEQUENCE 396 AA; 43577 MW; FCE4FB21972C3AD5 CRC64; E -> G (IN REF. 2).

Query Match 53.9%; Score 1171.5; DB 1; Length 396;
Best Local Similarity 59.9%; Pred. No. 4.8e-91;
Matches 233; Conservative 47; Mismatches 100; Indels 9; Gaps 4;

QY 14 LVLLLVPAANGCGPGR-VYSGRRRRPRKLVPLAYKQSPNPEKTLGASGRYEGKIA 72
DB 10 LCCLALALPAQ-SCGPGRPGVGRRRYARQQLVPLKYQFVPGVVERTLGASGPAEGRVA 68
QY 73 RSERFKELTPNYPNDITFKDENTGADRLTORCKDRNLSTAVNMQMPGVRLRYTEG 132
DB 69 RGSERFRLVYNYPDITFKDENSGADRLMTERCKERYNLAIYANMMPGVRLRYTEG 128
QY 133 WDEGHHSESLAYEGRAVDITSDRDNKYGLLARLAVAGFDMVYYESRAHYHCSYKS 192
DB 129 WDEGHHHODSLHYEGRALDITSDRDNKYGLLARLAVAGFDMVYYESRNHVSYKA 188
QY 193 EHSAAKTGCGCFPGAQVRLSGARVALSAVRBGDRVLAMGEDSPFTSDVLTFLDRPH 252
DB 189 DNSTLAVRAGCGFPGNATYRLWSGERKGLRELHRGDWVLAADASGRVPTPVLLFLDRDLO 248
QY 253 RLRAFOVETODPPRLALTPAHLFTADNHTPEARFRATFASHVOPGQVVLVAGVGL 312
DB 249 RRASTFVAETEMPRKLLITPMHLVFAARGPAPAGDPAPFARLRLRAGDSVLAPGDAL 308
QY 313 QPARVAAVSTHVALGAYAPLTKHGTLLVEDVVASCAAVADHHLAQLAFWPLRLFHSIAW 372
DB 309 RPARVAARVAREAVGVFAFLTAHGTLLVNDVLASCYAVLSEHQWHAFAFLRLHL-- 367
QY 373 GSWTPG-----EGYHWYQQLYKRLGLL 396
DB 367 GALLPGAVOPTGMWYSRLYRLAEELL 395

Search completed: June 5, 2000, 08:19:54
Job time: 158 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2000, 07:56:48 ; Search time 52.13 seconds
(without alignments)
2969.733 Million cell updates/sec

Title: US-08-900-220-8

Perfect score: 1191
Sequence: 1 ATGGCTCTCTGACCAATCT.....CGAGAGAGCTACTGGCTGA 1191

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 segs, 6492525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/5C.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/5D.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/6.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq:*
7: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	967.6	81.2	1190	2	US-08-176-427B-3 Sequence 3, Appl1
2	967.6	81.2	1190	3	US-08-356-060A-2 Sequence 2, Appl1
3	419.8	35.2	1313	2	US-08-176-427B-7 Sequence 7, Appl1
4	419.8	35.2	1313	3	US-08-356-060A-4 Sequence 4, Appl1
5	407	34.2	1715	6	PCT-US95-02315-1 Sequence 1, Appl1
6	390.6	32.8	1425	3	US-08-356-060A-6 Sequence 6, Appl1
7	389	32.7	1576	2	US-08-748-581-5 Sequence 5, Appl1
8	389	32.7	1576	2	US-08-748-591-10 Sequence 10, Appl1
9	381.4	32.0	1277	2	US-08-176-427B-1 Sequence 1, Appl1
10	381.4	32.0	1277	2	US-08-356-060A-5 Sequence 5, Appl1
11	355.8	29.9	1056	3	US-08-176-427B-5 Sequence 3, Appl1
12	355.8	29.9	1056	3	US-08-356-060A-3 Sequence 3, Appl1
13	335.6	28.2	939	3	US-08-356-060A-7 Sequence 9, Appl1
14	316.2	26.5	1256	2	US-08-176-427B-9 Sequence 9, Appl1
15	316.2	26.5	1256	3	US-08-356-060A-5 Sequence 5, Appl1
16	101.4	8.5	144	6	PCT-US95-15463-2 Sequence 2, Appl1
17	101.4	8.5	144	6	PCT-US95-15923-2 Sequence 2, Appl1
18	99.8	8.4	144	6	PCT-US95-15463-1 Sequence 1, Appl1
19	99.8	8.4	144	6	PCT-US95-15923-1 Sequence 1, Appl1
20	44.6	3.7	1642	1	US-08-232-015-2 Sequence 2, Appl1
21	43	3.6	1686	3	US-08-648-657-14 Sequence 14, Appl1
22	43	3.6	2504	1	US-08-073-384C-3 Sequence 3, Appl1
23	43	3.6	2504	1	US-08-254-359A-3 Sequence 3, Appl1
24	43	3.6	2504	2	US-08-483-043-3 Sequence 3, Appl1
25	43	3.6	2504	2	US-08-481-238-3 Sequence 3, Appl1
26	43	3.6	2504	3	US-08-471-066B-3 Sequence 3, Appl1
27	43	3.6	2504	3	US-08-484-956-3 Sequence 3, Appl1

C 28	43	3.6	2504	3	US-08-757-653-3 Sequence 3, Appl1
C 29	43	3.6	2504	3	US-08-599-491-3 Sequence 3, Appl1
C 30	43	3.6	2504	4	US-08-756-386-3 Sequence 3, Appl1
C 31	43	3.6	2504	4	US-08-823-516-3 Sequence 3, Appl1
C 32	43	3.6	2504	5	US-08-682-853A-3 Sequence 3, Appl1
C 33	43	3.6	2505	1	US-07-977-434-9 Sequence 9, Appl1
C 34	43	3.6	2505	2	US-08-458-819-9 Sequence 9, Appl1
C 35	43	3.6	2505	6	PCT-US91-07035-9 Sequence 9, Appl1
C 36	43	3.6	2640	1	US-08-384-490-30 Sequence 30, Appl1
C 37	43	3.6	2640	2	US-08-459-383-30 Sequence 30, Appl1
C 38	42.4	3.6	4257	3	US-08-680-473-1 Sequence 1, Appl1
C 39	42.4	3.6	9757	1	US-08-093-453B-1 Sequence 1, Appl1
C 40	42.4	3.6	9759	1	US-08-459-041A-1 Sequence 1, Appl1
C 41	42.4	3.6	12001	2	US-08-458-568A-11 Sequence 11, Appl1
C 42	41.6	3.5	20235	2	US-07-642-734C-3 Sequence 3, Appl1
C 43	41.6	3.5	30001	1	US-08-125-468-1 Sequence 1, Appl1
C 44	41.6	3.5	30001	3	US-08-474-933-1 Sequence 1, Appl1
C 45	41.4	3.5	1794	6	PCT-US95-14418-3 Sequence 3, Appl1

ALIGNMENTS

```
RESULT 1
US-08-176-427B-3
: Sequence 3, Application US/08176427B
: Patent No. 5789543
:
: GENERAL INFORMATION:
: APPLICANT: Ingham, Phillip W.
: APPLICANT: McMahon, Andrew P.
: APPLICANT: Tablin, Clifford J.
: TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
: TITLE OF INVENTION: Proteins and Uses Related Thereto
: NUMBER OF SEQUENCES: 33
: CORRESPONDENCE ADDRESS:
: ADDRESS: LAHIVE & COCKFIELD
: STREET: 60 State Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02109
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: ASCII(text)
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/176,427B
: FILING DATE: 30-DEC-1993
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Vincent, Matthew P.
: REGISTRATION NUMBER: 36,709
: REFERENCE/DOCKET NUMBER: HMI-006
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 227-7400
: TELEFAX: (617) 227-5941
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1190 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: both
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1191
:
: US-08-176-427B-3
:
: Query Match 81.2%; Score 967.6; DB 2; Length 1190;
: Best Local Similarity 88.3%; Pred. No. 5e-194;
: Matches 1051; Conservative 0; Mismatches 139; Indels 0;
```


Db	121	GTGGCTTGCTATACAAAGCAGTTTGTGGCCAGATGCCGACGGAGACCTTGGGGCGCAGT	181
Qy	181	GGGCGACCGAGGGGAGGGTGGCAGAGGGCTCCGAGCGCTTCGGGACCTGTGGTCCCAAC	240
Db	181	GGGGCGACCGAGGGGAGGGGTACAAAGGGGGTGGAGCGCTTCGGGACCTGTATCCCAAC	240
Qy	241	TACAAACCCGACATCACTTCAAGGATGAGGAGAAACAGTGGAGCCGACCGGCTGATGACC	300
Db	241	TACAAACCCCGACATATCTTCAAGGATGAGGAGAAACAGCGGGCGCACACCGGCTGATGACA	300
Qy	301	GAGCGTTGCAGAGAGAGGGTGAACGGCTTTGGCCATTGCCGTATGAAATATGTGGCCCGA	360
Db	301	GAGCGTTGCAGAGAGCGGGGTGAACGGCTCTAGCCATCGCGGTATGAACATGTGGCCCGA	360
Qy	361	GTGGCGCTACGAGTGAAGTGAAGGGCTGGGACAGAGACGGCCACCACGCTCAGGATTCAC	420
Db	361	GTAGCGCTACGTTGATGTGAAGGCTGGGACAGAGAGCGCCACACGACAGGATTCAC	420
Qy	421	CACATACAAAGCGCGTGTCTTGAGATCACTACGTCGACCGCGGACCGCAACAATGATGG	480
Db	421	CACATACAAAGCGCGTGTCTTGAGATCACTACGTCGACCGCGGACCGCAACAATGATGT	480
Qy	481	TTGCTGGCGCGCTCGCAGTGGAAAGCCGGCTTCGACTGGGTCTACTACGATCCCGCAAC	540
Db	481	TTGTTGGCGCGCGCTTACGCTGTGAAAGCCGGATTGCACTGGGTCTACTACGATCCCGCAAC	540
Qy	541	CAGTCCACAGTGTGGGTCAAGAGCTGATTAATCACTGAGGGGCTCGGGGGCGGGCGCTT	600
Db	541	CACATCCACGATGTGGGTCAAGAGCTGATTAATCACTGAGGGGCTCGGAGCGGAGCGCTT	600
Qy	601	CCGGGAATGCAACTGTGGCCCTGTGTGAGCGGCGAGCGGAAAGGGACTCGGGAATCGAC	660
Db	601	CCGGGAATGCAAGGTGGGTCAAGAGCTGATTAATCACTGAGGGGCTCGGAGCGGAGCGCTT	660
Qy	661	CGCGGAATGCAAGGTGGGTGGTGGGCGCCGATGCGTCAAGCGCGGGTGTGCCACCGCGTGTG	720
Db	661	CGTGGTACATGGGTGAGGGCGCGCTGTATGACAGCGCGCGGAGTGTATCCACCGCACTGTG	720
Qy	721	CTTCTTCGAGACGGGACCTTGACAGCGCGGGGCTTCAATTGTGGCTGTGAGACGAGTGG	780
Db	721	CTTCTTCGAGACGGGATCTTGACAGCGCGGGGCTTCAATTGTGGCTGTGAGACGAGTGG	780
Qy	781	CTTCCACGCAAACTGTGTCTACAGCGCCCTGGCACCTGTGTGGCGCTGAGAGGCCGCG	840
Db	781	CTTCCGCGCAAACTGTGTCTACACCCCTGGCACCTGTGTGTGGCTGAGAGGCCACG	840
Qy	841	CCGCGGACGCGACCTTTGACACCGGGTTCGCGCGCGGCTACGCGGTGAGGATCTCGGTG	900
Db	841	CTCTCTCAAGTGTACTTTGACACCGGTTCGCGCGCGGCTTACGTGTGTGGCATCTCGGTG	900
Qy	901	CTGGCGCGCGGGGATGTGCTTGGGCCACAGCGCGGTGGCCCTGTGTGGCGGGAGAGA	960
Db	901	CTGGCGTCCCGCGCGGGAGAGGGGCTCACAGCGCGCGGCTGTAGCCCGGTGGCGCGAGGA	960
Qy	961	GCGGTGGGCGTGTTCGCGCGCGCTCACCGCGCACAGGACGCTGTGTGAGATGATCTGTG	1020
Db	961	GCGGTGGGCGGTTCGCGACCGCTACCTGCGCACAGGAGACCTGTGTGTAACACATCTCTC	1020
Qy	1021	GCGCTTGTCTACGGGTTGTGAGAGATCACTAGGAGCGACCGCGCTTGTCCCGCTTG	1080
Db	1021	GCGCTTGTCTACGGGTTGTGAGAGATCACTAGGAGCGACCGCGCTTGTCCCGCTTGTG	1080
Qy	1081	AGACTGCTGCAAGCGCTTAGGGGGGCTGTCTCCCGCGGGGGCGCTCCAGCGCATGGCATG	1140
Db	1081	CGCGCTGCTGCAAGCGGCTTAGGGGGCTGTCTCCCTGGGGGCTGCAAGCGCATGGCATG	1140
Qy	1141	CATTGTACTCTCGGCTCTCTTACCGCTTAGCGGAGAGAGACTACTGGGCTG	1190
Db	1141	CATTGTACTCTCGGCTCTCTTACCGCTTAGCGGAGAGAGATATGGGCTG	1190

RESULT 3
US-08-176-427B-7

Sequence 7 Application US/08176427B
Patent No. 5789543

GENERAL INFORMATION:

APPLICANT: Ingham, Phillip W.
APPLICANT: Ingham, Andrew P.
APPLICANT: Tablin, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
TITLE OF INVENTION: proteins and Uses Related Thereto
NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/176,427B
FILING DATE: 30-DEC-1993
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMT-006

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 1313 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: linear

MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: CDS
LOCATION: 1..1314

US-08-176-427B-7

	Query Match	Best Local Similarity	35.2%	Score 419.8	DB 2	Length 1313
	Matches 701	Conservative	64.5%	Pred. No. 1e-79	Mismatches 362	Indels 24
					Gaps 4	
QY	41	CACCTTCGCGCGCTGCCACCCAGACGCTCGGGCCGGCGCGGGCGCGGTGTGGCCGCGCC	100			
Db	47	CGCTGCTGTGTGCCCCCGGGCTGGCGCTTGGGCCCGCGGAGGGGGTGTGGAAAGAGCGCG	106			
QY	101	CGTATGCGCGCAGACGCTCGTGGCGTACTCTACACGCAATTGTGCCCCGGCGTCCAG	160			
Db	107	AC-----CCCAAAAAGGTGACCCCTTTAGGCTACAAACAGATTATTTCCCAACGTAGCG	160			
QY	161	AGCGGACCCCTGGGGGCCAGTGGGCCAGCGGAGGGGAGGGTGGCAAGGGCTCCGAGGCGT	220			
Db	161	AGAGAACCCCTAGGGGCCAGCGGCGAGATGTGAAGGAAATCACAAGAAATCTCCGAGCAT	220			
QY	221	TCCGGACCTGTGTCCTCACTACAACCCCGCATCTTCACAGGATGAGGAAACAGT	280			
Db	221	TTAAGGACTACACCCCAATTACAACCCCGCATCTATTTAAGGATGAGGAAACAGG	280			
QY	281	GAGCGGACCGCTGTATGATCCGAGCGGTTCAGAGAGAGGGGTGAACGCTTTGGCCATTCGCG	340			
Db	281	GAGCGAGCCGGCTGTATGACTCAGAGGTGCAAGACAAATTTAAATCCCTTGGCCATCTCTG	340			
QY	341	TGATGAATCTTGCGCCGAGTGGCCCTACAGATGACTCAGAGGCGTGGGAGACAGAGGCGCC	400			
Db	341	TGATGAACCACTGGCGCTTGAGTGAAGCTGCACAGTATCCAGAGGCGTGGGATAGAGAGGCGCC	400			

QY 401 ACCAGCTCAGGATTCACCTCCACTACGAAAGCCGCTTGTGACATCACTACGTCTGACC 460
DB 401 ATCATCTCAGAGAGTCTCTACACTATGAGGCTGAGAGATGACATACACAGCTCCGACC 460
QY 461 GCGACCGCAACAAGTATGAGTGTGCGGCGCGCTCGACAGTGAACCGCGCTCGACTGGG 520
DB 461 GGGACCGCAACAAGTATGAGTGTGCGGCGCGCTCGACAGTGAACCGCGCTCGACTGGG 520
QY 521 TCTACTAGAGTCCCGCAACAAGTATGAGTGTGCGGCGCGCTCGACTGAATGATGAGGCG 580
DB 521 TCTACTAGAGTCCCGCAACAAGTATGAGTGTGCGGCGCGCTCGACTGAATGATGAGGCG 580
QY 581 TCCGGGGCGGCGCTGCTTCCGGGAATGCAACTGTGGGCTGTGAGACCGCGGCGGCA 640
DB 581 CCAAAATCCGGGCGGCTGCTTCCGGGAATGCAACTGTGGGCTGTGAGACCGCGGCGGCA 640
QY 641 AAGGCTCGGCAACTGCAACCGGAGACTGAGTGTGCGGCGCGCTCGACTGAGGCGGCGG 700
DB 641 AGCTGTGAGAGTCTAGTCTCCGGAGACCGGCTGTGCGGCGCTCGACTGAGGCGGCGGCG 700
QY 701 TGTGCGGCGGCGCTGCTGCTCTCTCTGAGACCGGAGACTTGTGAGGCGGCGGCTGATTTG 760
DB 701 TGTGCGGCGGCGCTGCTGCTCTCTCTGAGACCGGAGACTTGTGAGGCGGCGGCTGATTTG 760
QY 761 TGGCTGTGAGAGCCAGTGGGCTCCACGCAACTGTGCTGACGCGCTGAGCACTGTGTCT 820
DB 761 ACGTGTGAGAGCCAGTGGGCTCCACGCAACTGTGCTGACGCGCTGAGCACTGTGTCT 820
QY 821 TTTGCGGCTC-----GAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTGT 868
DB 821 TGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 880
QY 869 TCGGCGGCGGCGGCTGAGCGCTGAGGAGTCTGCTGAGGCGGCGGCGGCGGCGGCGG 925
DB 869 TCGGCGGCGGCGGCTGAGCGCTGAGGAGTCTGCTGAGGCGGCGGCGGCGGCGGCGG 925
QY 881 GCGTGTGCGGCGGCGGCGGCGGCTGTGAGTGTGCTGAGGAGTCTGAGGCGGCGGCGG 940
DB 881 GCGTGTGCGGCGGCGGCGGCGGCTGTGAGTGTGCTGAGGAGTCTGAGGCGGCGGCGG 940
QY 926 GAGCGGCGGCGGCTGAGGCGGCTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 982
DB 926 GAGCGGCGGCGGCTGAGGCGGCTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 982
QY 941 TCGGCGGCGGCGGCTGAGGCGGCTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1000
DB 941 TCGGCGGCGGCGGCTGAGGCGGCTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1000
QY 983 TCGGCGGCGGCGGCGGCGGCTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1042
DB 983 TCGGCGGCGGCGGCGGCGGCGGCTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1042
QY 1001 TCGGCGGCGGCGGCGGCGGCTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1060
DB 1001 TCGGCGGCGGCGGCGGCGGCTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1060
QY 1043 AGAGTCAAGTGTGAGGCGGCGGCTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1102
DB 1043 AGAGTCAAGTGTGAGGCGGCGGCTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1102
QY 1061 AGAGTCAAGTGTGAGGCGGCGGCTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1120
DB 1061 AGAGTCAAGTGTGAGGCGGCGGCTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1120
QY 1103 GCGGCT 1109
DB 1103 GCGGCT 1127

RESULT 4

US-08-356-060A-4
Sequence 4, Application US/08356060A
Patent No. 5844079

GENERAL INFORMATION:

APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tablin, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,060A
FILING DATE: 14-DEC-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/176,427
FILING DATE: 30-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMT-006CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ. ID NO. 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1313 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1314
US-08-356-060A-4

Query Match 35.28; Score 419.8; DB 3; Length 1313;
Best Local Similarity 64.58; Pred. No. 1e-79;
Matches 701; Conservative 0; Mismatches 362; Indels 24; Gaps 4;

QY 41 CACTTGTGCGGCTGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 100
DB 47 GCGTGTGCTGTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 106
QY 101 GCTATGCGGCGGCAAGCAGTCTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 160
DB 107 AC-----CCCAAAAGAGTGAACCGCTTTAGCTGCAAGAGAGAGTATTTCCAGAGTACCG 160
QY 161 AGCGGAGCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 220
DB 161 AGCGGAGCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 220
QY 221 TCCGGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 280
DB 221 TTAAGAGACTACCCCAATTTACACCCCGGACATCATTTTAAAGATGAGGAGAGAGAGAG 280
QY 281 GAGCGGAGCGGCTGATGAGCGGAGGCTTGAAGAGAGGCTGAAGCGCTTGGCCATTGCCG 340
DB 281 GAGCGGAGCGGCTGATGAGCGGAGGCTTGAAGAGAGGCTGAAGCGCTTGGCCATTGCCG 340
QY 341 TGAAGAGATGTGCGGCGGAGTGTGCTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAG 400
DB 341 TGAAGAGATGTGCGGCGGAGTGTGCTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAG 400
QY 401 ACCAGCTCAGAGTTCACCTCCACTACGAGAGGCGGCTGTTGAGATCACTACAGCTGAC 460
DB 401 ATCATTTAGAGAGTCTCTACACTATGAGGCTGAGAGAGTGTGAGATCACTACAGCTGAC 460
QY 461 GCGACCGCAACAAGTATGAGTGTGCGGCGCGCTCGAGTGAAGCGGCTGAGTGGG 520
DB 461 GGGACCGCAACAAGTATGAGTGTGCGGCGCGCTCGAGTGAAGCGGCTGAGTGGG 520
QY 521 TCTACTAGAGTCCCGCAACAAGTATGAGTGTGCGGCGCGCTCGAGTGAATGATGAGGCGG 580
DB 521 TCTACTAGAGTCCCGCAACAAGTATGAGTGTGCGGCGCGCTCGAGTGAATGATGAGGCGG 580
QY 581 TCCGGGGCGGCGGCTGCTTCCGGGAATGCAACTGTGGGCTGTGAGACCGCGGCGGCGG 640
DB 581 CCAAAATCCGGGCGGCTGCTTCCGGGAATGCAACTGTGGGCTGTGAGACCGCGGCGGCGG 640

```

0Y 641 AAGGCTCGGGAATGCAACCGGAGATGGGATTTTGCGGCCATGCTGTCAGGCCGG 700
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 641 AGCTGATGAGBACTTACGTCCCGAGACCGGCTCTGCGGCTACGACAGGCGCCGG 700
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
0Y 701 TGTGCCCCACGCGGCTGTCTCTTCTGAGACCGGGACTTGCAGCGCCGGCTTCATTG 760
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 701 TGCTATACAGGACATCTCTCACTCTCTGAGACCGGACGAGGCGCCAAAGAGTCTCT 760
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
0Y 761 TGGCTGTGAGACCAAGTGCCCTCCACGCAACTGTGCTACGCGCTTGCACTGTGT 820
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 761 ACGTATGAGAGACGCTGAGCGCGCGAGACGCGCTGCTACCGCCGCGACACTGTCT 820
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
0Y 821 TTGCCGCTC-----GAGGCGCGCGCCGCGCGACGACACTTTCACCGGTGT 868
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 821 TCGTGCGCGCGGACACAGCAAGCTCGGGGCCACGCGCCGCGGACAGCGCGCTCTTTCACAGC 880
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
0Y 869 TCGCGCGCGCGCTACGCGCTGGGAGACTCGGTGCTGCGCGCGCGCGGGA--TGCCTTC 925
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 881 GCGTCCGCGCGCGCGGACGCGCTGTACGTGTGTGAGACGCGCGGAGCGCGCGCTGC 940
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
0Y 926 GGGCAGCGCGCTGCGCCCGTGTGCG--CGGGAGAGAAAGCGTGGGCGTGTTCGCGCGC 982
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 941 TCGCGCGCGGTTGACAGCGTGAAGCTGCGGAGAGAGAGAGCGCGCGCGCTACGCGCGC 1000
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
0Y 983 TCACGCGCGACGAGCGGTGCTGAGAGAGATGCTGCGCTCTTGTCTACGCGGTTCTG 1042
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1001 TCACGCGCGACGCGACACATTCATTCACACCGGCTGCTGCGCTGTACGCTGTATCG 1060
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
0Y 1043 AGAGTCACCAAGTGGCGGACCGCGCTTTTGCCCGCTTGAGACTGCTGACGCGCTAGGG 1102
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1061 AGGACACAGCTGGCGACACCGGCGCTTGCGCGCTTCCGCTGCGCGACGCGGCTGTG 1120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
0Y 1103 CGCTGCT 1109
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1121 CCGCGCT 1127
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
PCT-US95-02315-1
: Sequence 1, Application PC/TUS9502315
: GENERAL INFORMATION:
: APPLICANT: Jessell, Thomas M.
: APPLICANT: Dodd, Jane
: APPLICANT: Roelink, Henk
: APPLICANT: Edlund, Thomas
: TITLE OF INVENTION: DNA ENCODING A VERTEBRATE HOMOLOG OF
: TITLE OF INVENTION: HEDGEHOG, VHH-1, EXPRESSED BY THE NOTOCHORD, AND USES
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Cooper & Dunham
: STREET: 1185 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/02315
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: John P. White
: REGISTRATION NUMBER: 28,678
: REFERENCE/DOCKET NUMBER: 45375-A-PCT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 391-0400
: TELEFAX: (212) 391-0525

```

```

1 INFORMATION FOR SEQ ID NO: 1
2
3 SEQUENCE CHARACTERISTICS:
4
5     LENGTH: 1715 base pairs
6
7     TYPE: nucleic acid
8
9     STRANDEDNESS: single
10
11     TOPOLOGY: linear
12
13     MOLECULE TYPE: cDNA
14
15     FEATURE:
16
17     NAME/KEY: CDS
18     LOCATION: 315..1628
19
20     CDS-UTS5-02315-1

```

Query Match	34.2%;	Score 407;	DB 6;	Length 1715;
Best Local Similarity	63.8%;	Pred. No. 4.8e-77;		
Matches 693;	Conservative	0;	Mismatches 370;	Indels 24;
			Gaps	4;

[illegible]


```

: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Vincent, Matthew P.
: REGISTRATION NUMBER: 36,709
: REFERENCE/DOCKET NUMBER: HMT-006
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 227-7400
: TELEFAX: (617) 227-5941
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1277 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: both
: TOPOLOGY: linear
: MOLECULE TYPE: cdna
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1277
:
: US-08-176-427B-1

```

Query Match	32.08;	Score 381.4;	DB 2;	Length 1277;
Best Local Similarity	62.08;	Pred. No. 1e-71;		
Matches 692;	Conservative 0;	Mismatches 386;	Indels 39;	Gaps 4;

QY	109	CGCAGAGGCTCGTGGCCGCTCTCTACAGCAATTTGGCCCGGCGTGCAGAGCCGAC	168
Db	115	CCCAAAAGCTGACCCCGCTTAGCTTATAGACATTATTCCAATGTG bb CGAGAAACC	174
QY	169	CTGGCGCCCACTGGGCGCAGCGGAGGGGAGGGTGTGCAGAGGGGCTCCAGCCCTCCGGAC	228
Db	175	CTAAGGGGCACTGGAAAGTATGAAAGGGAAGTACAGAAACTCCGAGAGTTTAAAGAA	234
QY	229	CTCGTCCCAACTCAACACCCCGACATCATCTTCAAGATGAGAGAAACAGTGAACCGAC	288
Db	235	CTAACCCCAATTACACCCCTGACATTATTTTAAAGATGAAGAGAACACGGGAGCTGAC	294
QY	289	CGCCTGATGACCGAGCGCTGTGAAGAGAGGTGAACCTTTGGCGATTGGCCGTATATAC	348
Db	295	AGACTGATGACTACGGCTGTGAAGAGCAACGCTGAATCCCTGGGGAGTCTCGGTATATAC	354
QY	349	ATGTGGCCCGGAGAGCGGCTTACAGAGTGCATGAGGAGCTGGGACGAGAGCGGCACACAGCT	408
Db	355	CAGTGGCCCCGGGGGTGAAGCTCGGGGTGACCGAGGGCTTGGAGCGAGATGGCATCACTCC	414
QY	409	CAGATTTACATCCACTAGAGAGCGCGTGTGGACATCACTACGTGTGAACCGGACCGC	468
Db	415	GAGGAATGCTGACACTACGAGAGGTGGCGCGTGGACATACACAGCTGGATCGGGACCGC	474
QY	469	AACAAGTATGGGTGCTGGCGCGCTCGCACTGGAAACCGGCTTCGACTGGGTCTACTAC	528
Db	475	AGCAAGTACGGAAAGTGTGGCCCGCTCGCCCTCGAGCGCGGCTTCGACTGGGTCTACTAC	534
QY	529	GAGTCCCGCAACACGTCACAGTCTCGGTCAAACTATTACTACATGGCGGTCCGGGCG	588
Db	535	GAGTCCAAAGCGCACATCACTGCTCCGTCAAAACAGAAACTAGTGGACGCGAAATCA	594
QY	589	GCGCGCTGCTTCCGGGAAATGACAATGTGGCCTGTGGAGCGGAGCGGAAAGGCGCT	648
Db	595	GGAGGCTGCTTCCCTGGCTGACCCACAGTGTGACTTGGAGCATGGAAGGACCAACAGCTGTG	654
QY	649	CGGGAATGCACCGCGAGACTGGGTTTTGGCGGCGCATGCGTCAGAGCGCGGGTGTGCC	708
Db	655	AAGGACCTGAGCCCTGGGGAGCCGGTGTGCTGTGACCGGAGCGCGCGGCTCTCTAC	714
QY	709	ACGCGGATGCTCTTCTTCTTGACCGGGACTTGCAGCGCGCGGGCTTCATTGTGTGCTGTG	768
Db	715	AGTGACTTCTCACTCTCTCTCGACCGGATGGACAGCTCCGAAAGCTCTTCTACGTCATC	774
QY	769	GAGACCGAGTGGCGTCCACGCAAACTGTGTGTCAGCGCCTGACACCTGGATTGGCGCT	828
Db	775	GAGACGGCGGACGCCCGGGCGCGGCTGCTACTGACGGCGGCCCACTGCTTTGTGTGCC	834

QY	829	CGAGGGCGGGCGCCGGGCGCAGGCG-ACCTTACCGAGGTTCGGGCGC-----	877
Db	835	CCCCAGCACACCAAGTCGAGAGGCCACAGAGGTCCACCAAGTGGCCAGGCGCTCTTCCGCACAG	894
QY	877	-----CGGCTACGCGCGCTGGGAGGACTCGGTGTCTGGGCGCCCGCGGGGAGATCGCTTGGGCGCA	930
Db	895	AAGGTGAAGCGCTGGCCAAAGTGTCTATGTGTGTGGGAGAGGGGGGCGACAGAGCTGTGTCGG	954
QY	931	GGCGGCGCTGGCCCGGTGTGGC--GCGGAGAGAACCCGTGGCGTGTGGCGCGCTACAC	987
Db	955	GCGCTCTTCACACAGCGTCTCATTTGTCGGGAGAGAGGCGTCCGGAGCCTACGAGCCCACTACAC	1014
QY	988	GCGCAGCGGAGCGCTGCGGTGGAGAACATATGCTCGAGCCTCTGTGTACGCGGTGTGGAGAGT	1047
Db	1015	GCCCCAGGACCAATCCTCATCAACCGGGGTGTGGCTCTCTGTATAGCGCGTCATCGAGAG	1074
QY	1048	CACCAATGGGAGGAGACGCGCGCTTTTGGCCCTTGAGACTGCTACAGCGCGTAGGAGCGCGTG	1107
Db	1075	CACAGTTGGGGCCCATTTGGGCGCTTGGGACCATTTCCGCTTGCGTTCAGAGGGGTGCTGTGGCGCC	1134
QY	1108	C-----TCCCCGGCGGGCGCTGCACCGCAGTGGCATGTGATTTGGTAC	1149
Db	1135	CTCTGCCAGATGGGGCCATCCCTACTTGGCGGCACCCACACCACTGCGATCCATCATGTGGTAC	1194
QY	1150	TCTGGCTCCTTACAGCGCTTAGGAGAGAGCTACTGG	1186
Db	1195	TCAAGGCTCCTTACCGCATCGCATGGCAGCTGGGTGCTGG	1231

RESULT 10
 US-08-356-060A-1
 Sequence 1, Application US/08356060A
 Patent No. 5844079
 GENERAL INFORMATION:
 APPLICANT: Ingham, Phillip W.
 APPLICANT: McMahon, Andrew P.
 APPLICANT: Tablin, Clifford J.
 TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
 TITLE OF INVENTION: Proteins and Uses Related Thereto
 NUMBER OF SEQUENCES: 47
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: LAHIVE & COCKFIELD
 STREET: 60 State Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: ASCII(text)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/356,060A
 FILING DATE: 14-DEC-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/176,427
 FILING DATE: 30-DEC-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Vincent, Matthew P.
 REGISTRATION NUMBER: 36,709
 REFERENCE/DOCKET NUMBER: HMT-006CP
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 227-7400
 TELEFAX: (617) 227-5941
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1277 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: both
 TOPOLOGY: linear
 MOLECULE TYPE: CDNA

TITLE OF INVENTION: Proteins and Uses Related Thereto
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/176,427B
FILING DATE: 30-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMI-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1256 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1257
US-08-176-427B-9

Query Match 26.5%; Score 316.2; DB 2; Length 1256;
Best Local Similarity 58.4%; Pred. No. 4,1e-58;
Matches 571; Conservative 0; Mismatches 403; Indels 3; Gaps 1;

112 AAGCAGCTGCGCCCTTCTACAGCAATTTGTCGCCGCGCTCCAGAGCGGACCTG 171
109 AAGAGCTGACACCTCTGCGCTACAGCAATCTATCAATGCGGAGAAAGCTTA 168
172 GCGCCAGTGGGCGCCAGGAGGAGGCTGCGCAAGGCGTCCGACCGCTCCGGAGCTC 231
169 GGGCCAGCGGAGATACAGGAGGCAATACGCGCAATTCGAGAGATTAAAGAACTT 228
232 GTGCCAATCAACCCGACATCATCTTCAAGATGAGAGAGACAGGAGCGGACCGC 291
229 ACTCAATTTACATCCCGACATTTCTTTAAGATGAGAGAGAAACGAGGAGCAGG 288
292 CTGATGACCGAGCGTTGGAAGAGAGGTTGAACGCTTGGCCATTGCCGTATGAATG 351
289 CTATGACACAGAGATGCAAAAGCAAGCTGAACCTGCGTGGCATCTCTATATGAACGC 348
352 TGGCCGAGAGTCGCGCTACGAGTCTGAGGCGTGGAGAGAGAGCGGACCGACCTCAG 411
349 TGGCCAGGAGTTAACTGCTGCTGACAGAGGCGTGGATGAGAGAGGTCACCATTTTGA 408
412 GATTCATCTCATAGAGAGCGCGCTTGGACATCATAGTCTGACCGGAGCGGCAAC 471
409 GATATCTCTCATAGAGAGAGAGAGAGCTGTTGATATTACACTCTGACCGAGACAGAGC 468
472 AAGTATGGTTCGCGCGCTCTCCAGTGAAGCGGCGCTGAGTGGTCTACTACGAG 531
469 AATATGGGAGACCTGTCTGCTACCTGAGAGGCTGATTTGATGGGTCTATATACAG 528
532 TCCGCAACACAGTCTGAGTCTGCTCAAGCTGATTAAGTCTACTGCGGCTCCGGGCGGC 591
529 TCCAAAGCCACATTCATGCTGTCTCAAGAGAAATTCGTTGCTGCGAAATCTGGG 588

592 GCGTCCTTCGCGGAAATGCAACTGTGTGCGCTGTGAGCGGCGAGGAGAAAGGCTGCGG 651
589 GAGCTTTTCCAGAGTTTCGCTGTGTGTGCTCCAGAGCGGAGGAGCAAGAGCGGCTGAG 648
652 GACTGCAACCGGAGAGTGGGTTTGGCGGCGGACATGCGGTAGGCGGCGGCTGCCAGC 711
649 GACCTGAACCCCGGAGACAGAGTGTGCGGCGGAGACAGCGGCGGAAACCTGTGTTCAGC 708
712 CCGGTGCTCTCTTCTGACCGGAGCTTGACAGCGCGGCGGCTTATTTGTGCTGTGAG 771
709 GACTTCATCATGTTTACAGACCGAGACTTCACAGCGGAGAGCTGTGTTTACGTATGAA 768
772 ACCGAGTGGCTTCACGCAACTGTTGCTACAGCGGCTGCGACCTGTGTTTCCGCTCGA 831
769 ACGCAAGAACCGTGAAGAAAGTACACCTCAACCGCGGCTACACTCTTTTGTCTGAC 828
832 GGGCGGCGCGCGCGGCGGAGCTTTCACCGGCTGTGCGCGCGCGGCTAGCGGCTGG 891
829 AACTCAACGGAAGATCTCCACACATGACCCCGGCTATGCCAGAGTGCAGAGCCGGA 888
892 GACTGCGTGGCGCGCGCGCGGCGG---GGATGCGCTTTCGCGGAGCGGCTGCGCTGTG 948
889 CAAAGGTGATGTTGTTGATGATAGCGGTACGCTTAATCTGTATGTGACAGGAGATA 948
949 GCGCGGAGAGAGCGGTGGCGGTGTTGCGCGCTACCGCGGACGAGAGCGTGTGAGTG 1008
949 TACACGAGAGAGACAGCGGCGCTGCTCCACACAGTATGACATGAGAACATTTTGCTG 1008
1009 AAGATGCTGCGCTCTGTTGACGCGGTTCTGAGAGTCAACAGTGGCGGACCGCGCT 1068
1009 GACAGAAATACGCGCTGCTTACCGCGCTTAATAGAGACACAGGCGCTTGGCATTTGGCC 1068
1069 TTTGCGCGCTTGAGACT 1085
1069 TTGCGCGCGCGCGGCT 1085

RESULT 15

US-08-356-060A-5
Sequence 5, Application US/08356060A

Patent No. 5844079

GENERAL INFORMATION:

APPLICANT: Ingham, Phillip W.

APPLICANT: McMahon, Andrew P.

APPLICANT: Tabin, Clifford J.

TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing

NUMBER OF SEQUENCES: 47

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD

STREET: 60 State Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII(text)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/356,060A

FILING DATE: 14-DEC-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/176,427

FILING DATE: 30-DEC-1993

ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.

REGISTRATION NUMBER: 36,709

REFERENCE/DOCKET NUMBER: HMI-006CP

TELECOMMUNICATION INFORMATION:

Thu Jun 8 15:53:43 2000

us-08-900-220-8.mli

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2000, 07:58:23 ; Search time 74.45 Seconds
(without alignments)
4002.401 Million cell updates/sec

Title: US-08-900-220-8

Perfect score: 1191
Sequence: 1 ATGGCTCTCCTGACCACTCTGAGAGAGCTACTGGAGCTGA 1191

Scoring table: IDENTITY_NDC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4190	99.9	1190	X07270	Human Deserit hedge
2	4190	99.9	1190	X07270	Human Deserit hedge
3	4190	99.9	1190	X07270	Human Deserit hedge
4	4190	99.9	1190	X07270	Human Deserit hedge
5	4190	99.9	1190	X07270	Human Deserit hedge
6	4190	99.9	1190	X07270	Human Deserit hedge
7	4190	99.9	1190	X07270	Human Deserit hedge
8	4190	99.9	1190	X07270	Human Deserit hedge
9	4190	99.9	1190	X07270	Human Deserit hedge
10	4190	99.9	1190	X07270	Human Deserit hedge
11	4190	99.9	1190	X07270	Human Deserit hedge
12	4190	99.9	1190	X07270	Human Deserit hedge
13	4190	99.9	1190	X07270	Human Deserit hedge
14	4190	99.9	1190	X07270	Human Deserit hedge
15	4190	99.9	1190	X07270	Human Deserit hedge
16	4190	99.9	1190	X07270	Human Deserit hedge
17	4190	99.9	1190	X07270	Human Deserit hedge
18	4190	99.9	1190	X07270	Human Deserit hedge
19	4190	99.9	1190	X07270	Human Deserit hedge
20	4190	99.9	1190	X07270	Human Deserit hedge
21	4190	99.9	1190	X07270	Human Deserit hedge
22	4190	99.9	1190	X07270	Human Deserit hedge
23	4190	99.9	1190	X07270	Human Deserit hedge
24	4190	99.9	1190	X07270	Human Deserit hedge
25	4190	99.9	1190	X07270	Human Deserit hedge
26	4190	99.9	1190	X07270	Human Deserit hedge
27	4190	99.9	1190	X07270	Human Deserit hedge
28	4190	99.9	1190	X07270	Human Deserit hedge
29	4190	99.9	1190	X07270	Human Deserit hedge
30	4190	99.9	1190	X07270	Human Deserit hedge
31	4190	99.9	1190	X07270	Human Deserit hedge
32	4190	99.9	1190	X07270	Human Deserit hedge
33	4190	99.9	1190	X07270	Human Deserit hedge
34	4190	99.9	1190	X07270	Human Deserit hedge

35	335.6	28.2	939	1	Q91641	Human Indian hedge
36	316.4	26.6	1251	1	X16189	Zebrafish Tth hedy
37	316.4	26.6	1251	1	X07277	Zebrafish tlgie-w
38	316.4	26.6	1251	1	X25106	Zebrafish tlgie-w
39	316.2	26.5	1256	1	Q91638	Zebrafish sonic he
40	316.2	26.5	1256	1	X16186	Zebrafish Shn hedy
41	316.2	26.5	1256	1	X07275	Zebrafish Sonic he
42	316.2	26.5	1256	1	X25102	Zebrafish Sonic he
43	299.8	25.2	1416	1	X16190	Drosophila HH hedy
44	299.8	25.2	1416	1	X07278	Drosophila hedgeho
45	299.8	25.2	1416	1	X25107	Drosophila hedgeho

ALIGNMENTS

RESULT	ID	Score	Query Match	Length	ID	Description
1	X07270	335.6	28.2	939	1	Q91641
2	X07270	316.4	26.6	1251	1	X16189
3	X07270	316.4	26.6	1251	1	X07277
4	X07270	316.4	26.6	1251	1	X25106
5	X07270	316.2	26.5	1256	1	Q91638
6	X07270	316.2	26.5	1256	1	X16186
7	X07270	316.2	26.5	1256	1	X07275
8	X07270	316.2	26.5	1256	1	X25102
9	X07270	299.8	25.2	1416	1	X16190
10	X07270	299.8	25.2	1416	1	X07278
11	X07270	299.8	25.2	1416	1	X25107
12	X07270	299.8	25.2	1416	1	X25107
13	X07270	299.8	25.2	1416	1	X25107
14	X07270	299.8	25.2	1416	1	X25107
15	X07270	299.8	25.2	1416	1	X25107
16	X07270	299.8	25.2	1416	1	X25107
17	X07270	299.8	25.2	1416	1	X25107
18	X07270	299.8	25.2	1416	1	X25107
19	X07270	299.8	25.2	1416	1	X25107
20	X07270	299.8	25.2	1416	1	X25107
21	X07270	299.8	25.2	1416	1	X25107
22	X07270	299.8	25.2	1416	1	X25107
23	X07270	299.8	25.2	1416	1	X25107
24	X07270	299.8	25.2	1416	1	X25107
25	X07270	299.8	25.2	1416	1	X25107
26	X07270	299.8	25.2	1416	1	X25107
27	X07270	299.8	25.2	1416	1	X25107
28	X07270	299.8	25.2	1416	1	X25107
29	X07270	299.8	25.2	1416	1	X25107
30	X07270	299.8	25.2	1416	1	X25107
31	X07270	299.8	25.2	1416	1	X25107
32	X07270	299.8	25.2	1416	1	X25107
33	X07270	299.8	25.2	1416	1	X25107
34	X07270	299.8	25.2	1416	1	X25107
35	X07270	299.8	25.2	1416	1	X25107

QY	61	CAGAGCTGGGGGCGGGGCGGGGGCGGGTGTGGCGGGGCGCCGCTATGCGGGCAAGGAGCTC	120
Dp	61	CAGAGCTGGGGGCGGGGCGGGGGCGGGTGTGGCGGGGCGCCGCTATGCGGGCAAGGAGCTC	120
QY	121	GTGGCGGTACTCTCAAGCAATTTGTGTCCCGGGGTGCCAGAGCGGACCCCTGGGGCCGAGT	180
Dp	121	GTGGCGGTACTCTCAAGCAATTTGTGTCCCGGGGTGCCAGAGCGGACCCCTGGGGCCGAGT	180
QY	181	GGGCGACGGAGGGGGGGGTGGCAGAGGGGCTCCGAGCGCTTCCGGGACTGTGCCCAAC	240
Dp	181	GGGCGACGGAGGGGGGGGTGGCAGAGGGGCTCCGAGCGCTTCCGGGACTGTGCCCAAC	240
QY	241	TACAAACCCCGACATCATCTTCAAGGATGAGGAGAACAGTGGAGCGGACCGCGCTGATGACC	300
Dp	241	TACAAACCCCGACATCATCTTCAAGGATGAGGAGAACAGTGGAGCGGACCGCGCTGATGACC	300
QY	301	GAGCGTTGCAAGGAGAGGGGTGAACCGCTTTGGCCATTGGCCGTGATGAACATGTGGCCGGA	360
Dp	301	GAGCGTTGCAAGGAGAGGGGTGAACCGCTTTGGCCATTGGCCGTGATGAACATGTGGCCGGA	360
QY	361	GTGGCGCTACGAGTGAAGAGGCTGGGACGAGAGCGGCCACCGCTCAGGATTCACCTC	420
Dp	361	GTGGCGCTACGAGTGAAGAGGCTGGGACGAGAGCGGCCACCGCTCAGGATTCACCTC	420
QY	421	CACATCAGAAAGGCGGTGGTTGSAATCATTACATAGCTGACCGGGACCGCAACAAAGTATGG	480
Dp	421	CACATCAGAAAGGCGGTGGTTGSAATCATTACATAGCTGACCGGGACCGCAACAAAGTATGG	480
QY	481	TTGCTGGCGGGCCTCGCAGTGTGAAGCGCGCTTGCATGGGTACTACAGATGTCCCGCAAC	540
Dp	481	TTGCTGGCGGGCCTCGCAGTGTGAAGCGCGCTTGCATGGGTACTACAGATGTCCCGCAAC	540
QY	541	CACGTCCAGTGTGGGTCAAAAGCTGATTAATCAGTGGCGGTCCGGGCGGGCTGTCTTT	600
Dp	541	CACGTCCAGTGTGGGTCAAAAGCTGATTAATCAGTGGCGGTCCGGGCGGGCTGTCTTT	600
QY	601	CCGGGAAATCACTGTGGGCTGTGGAGCGGGGAGCGGAAAGGCTGGGGAACTGCAC	660
Dp	601	CCGGGAAATCACTGTGGGCTGTGGAGCGGGGAGCGGAAAGGCTGGGGAACTGCAC	660
QY	661	CGCGGAGACTGTGGTTTGGGCGGCGGATGCGTCAGCGCGGGTGGTGCCACGCGGCTGTG	720
Dp	661	CGCGGAGACTGTGGTTTGGGCGGCGGATGCGTCAGCGCGGGTGGTGCCACGCGGCTGTG	720
QY	721	CTCTCTGGAACGGGACTTGCACAGCGCGGGGCTTCAATTTGTGGCTGTGGAGACGAGTGG	780
Dp	721	CTCTCTGGAACGGGACTTGCACAGCGCGGGGCTTCAATTTGTGGCTGTGGAGACGAGTGG	780
QY	781	CCCTCCAGGCAAACTGTGTGTCACAGCCCTGGCACTGTGTGGCTGTGAAGGCGCGCG	840
Dp	781	CCCTCCAGGCAAACTGTGTGTCACAGCCCTGGCACTGTGTGGCTGTGAAGGCGCGCG	840
QY	841	CCCGCGGACGAGGCACTTGTGCACCGGGTTTCGCGCGCGGCTACGCGCTGGGACTGTGGTG	900
Dp	841	CCCGCGGACGAGGCACTTGTGCACCGGGTTTCGCGCGCGGCTACGCGCTGGGACTGTGGTG	900
QY	901	CTGGCGCGCGGGGGGATGCGCTTGGGCCACGCGCGGCTGGCCCGCTGTGTGGCGGGAGAA	960
Dp	901	CTGGCGCGCGGGGGGATGCGCTTGGGCCACGCGCGGCTGGCCCGCTGTGTGGCGGGAGAA	960
QY	961	GCGCTGGGGCTGTTCGCGCGGCTCACCGGCAAGGAGACCTGTGTGAACATGTCTCTG	1020
Dp	961	GCGCTGGGGCTGTTCGCGCGGCTCACCGGCAAGGAGACCTGTGTGAACATGTCTCTG	1020
QY	1021	GCGCTTGTACAGGGGTTTGGAGAGTCAACAGTGGGGCACCGGGCTTTGCGCCCTTG	1080
Dp	1021	GCGCTTGTACAGGGGTTTGGAGAGTCAACAGTGGGGCACCGGGCTTTGCGCCCTTG	1080
QY	1081	AGACTGTGCACGGGCGTAAAGGGGGGCTGTCCCGGGGGGGCGTCCAGCCGACATGTGCATG	1140
Dp	1081	AGACTGTGCACGGGCGTAAAGGGGGGCTGTCCCGGGGGGGCGTCCAGCCGACATGTGCATG	1140
QY	1141	CATTGTACTCTGGCTCTCTACCGCTTAAGGAGAGACTACTGGGCTG	1190

Query Match	Best Local Similarity	Score	DB	Length
Matches 1190: Conservative 0; Mismatches 0; Indels 0; Gaps 0;	99.9%;	1190;	DB 1;	Length 1190;
0Y 1 ATGGCTCTCTGACCAATCTACTACTGCGCTTGTGCTCTTGGCACTTCTGGCGCTGCCAGCC 60	1 ATGGCTCTCTGACCAATCTACTACTGCGCTTGTGCTCTTGGCACTTCTGGCGCTGCCAGCC 60			
Db 1 ATGGCTCTCTGACCAATCTACTACTGCGCTTGTGCTCTTGGCACTTCTGGCGCTGCCAGCC 60				
0Y 61 CAGAGCTCGGGGCGGGGCGGGGCGGGTGTGGCCGGGCGCGGTATGCGCGCAAGAGGCTC 120	61 CAGAGCTCGGGGCGGGGCGGGGCGGGTGTGGCCGGGCGCGGTATGCGCGCAAGAGGCTC 120			
Db 61 CAGAGCTCGGGGCGGGGCGGGGCGGGTGTGGCCGGGCGCGGTATGCGCGCAAGAGGCTC 120				
0Y 121 GTGCGCTACTCTTACAAACAATTTGTGCGCGCGGTGCGAGAGCGAGCCCTGGGCGCCAGT 180	121 GTGCGCTACTCTTACAAACAATTTGTGCGCGCGGTGCGAGAGCGAGCCCTGGGCGCCAGT 180			
Db 121 GTGCGCTACTCTTACAAACAATTTGTGCGCGCGGTGCGAGAGCGAGCCCTGGGCGCCAGT 180				
0Y 181 GGGCGCAGCGAGAGGGAGGTGGCAAGGGGCTTCCAGGCGCTTCCGGGACCTGTGCGCCAAAC 240	181 GGGCGCAGCGAGAGGGAGGTGGCAAGGGGCTTCCAGGCGCTTCCGGGACCTGTGCGCCAAAC 240			
Db 181 GGGCGCAGCGAGAGGGAGGTGGCAAGGGGCTTCCAGGCGCTTCCGGGACCTGTGCGCCAAAC 240				
0Y 241 TACAAACCCGCAATCATCTTCAAGATAGAGGAGAACAGTGGAGCGCGACCGCTGATGACC 300	241 TACAAACCCGCAATCATCTTCAAGATAGAGGAGAACAGTGGAGCGCGACCGCTGATGACC 300			
Db 241 TACAAACCCGCAATCATCTTCAAGATAGAGGAGAACAGTGGAGCGCGACCGCTGATGACC 300				

```

QY 301 GAGGCTTGCAGAGAGAGGCTGAAGCCTTTGGCCATTTGCCGTGATGACATGTGCCCGGA 360
DB 301 GAGGCTTGCAGAGAGAGGCTGAAGCCTTTGGCCATTTGCCGTGATGACATGTGCCCGGA 360
QY 361 GTGGCCCTACAGAGTACTGAGAGGCTGGGACAGAGAGGCGCACAGGCTCAGATTACTC 420
DB 361 GTGGCCCTACAGAGTACTGAGAGGCTGGGACAGAGAGGCGCACAGGCTCAGATTACTC 420
QY 421 CACTACGAAGGCGGTGCTTTGGACATCACTAGTGTGACCGCGACCGCAACAGATGGG 480
DB 421 CACTACGAAGGCGGTGCTTTGGACATCACTAGTGTGACCGCGACCGCAACAGATGGG 480
QY 481 TTGCTGGCGCCCTTCGACAGTGAAGCCGGCTTCGACTGGGTCTACTACGATCCCGAAC 540
DB 481 TTGCTGGCGCCCTTCGACAGTGAAGCCGGCTTCGACTGGGTCTACTACGATCCCGAAC 540
QY 541 CAGGTCAGTGTGGTCAAGAGTGAATCACTACGAGGCGGTCCGGGGGGGGGGCTT 600
DB 541 CAGGTCAGTGTGGTCAAGAGTGAATCACTACGAGGCGGTCCGGGGGGGGGGCTT 600
QY 601 CCGGGAATGCACTGTGCGCTGTGAGCGCGGACGCGGAAGGGCTCGGGAATCGAC 660
DB 601 CCGGGAATGCACTGTGCGCTGTGAGCGCGGACGCGGAAGGGCTCGGGAATCGAC 660
QY 661 CGCGGAGACTGGGTTTGGCGCGCGAGTCGTACAGCGCGGGGTGGTCCACCGCGTCTG 720
DB 661 CGCGGAGACTGGGTTTGGCGCGCGAGTCGTACAGCGCGGGGTGGTCCACCGCGTCTG 720
QY 721 CTCTCCGAGCGGGGACTTCAGAGCGCGGGCTTCATTTGGGCTGTGGAACCGAGTGG 780
DB 721 CTCTCCGAGCGGGGACTTCAGAGCGCGGGCTTCATTTGGGCTGTGGAACCGAGTGG 780
QY 781 CCTCCAGCAAACTGTGCTACAGCGCCCTGAGCACTGTGTTCGCTCGAGAGCGGCG 840
DB 781 CCTCCAGCAAACTGTGCTACAGCGCCCTGAGCACTGTGTTCGCTCGAGAGCGGCG 840
QY 841 CCGCGGCGGAGCTTTGACACCGGTGTTCGCGCGCGCTACGCGCTGGGACTCGGTG 900
DB 841 CCGCGGCGGAGCTTTGACACCGGTGTTCGCGCGCGCTACGCGCTGGGACTCGGTG 900
QY 901 CTGCGCGCGCGGAGTGGCTTCGCGCGAGCGGGCGGTGGGGGGGGGGAGAA 960
DB 901 CTGCGCGCGCGGAGTGGCTTCGCGCGAGCGGGCGGTGGGGGGGGGGAGAA 960
QY 961 GCGCTGGCGGTTCGCGCGCGCTACCGCGAGGAGCTGTGTGTAAGATGCTCTG 1020
DB 961 GCGCTGGCGGTTCGCGCGCGCTACCGCGAGGAGCTGTGTGTAAGATGCTCTG 1020
QY 1021 GCGCTGTGCTACGCGGTTCGAGAGTACCAAGTGGGCGACCGCGCTTTGGCCCTTG 1080
DB 1021 GCGCTGTGCTACGCGGTTCGAGAGTACCAAGTGGGCGACCGCGCTTTGGCCCTTG 1080
QY 1081 AGACTGGCGAGCGGTTCGAGAGTACCAAGTGGGCGACCGCGCTTTGGCCCTTG 1140
DB 1081 AGACTGGCGAGCGGTTCGAGAGTACCAAGTGGGCGACCGCGCTTTGGCCCTTG 1140
QY 1141 CATGTGACTCTCGGCTCTTACCGCTTACGAGAGAGTACTGGGCTG 1190
DB 1141 CATGTGACTCTCGGCTCTTACCGCTTACGAGAGAGTACTGGGCTG 1190

```

```

RESULT 3
ID V62395 standard: cDNA: 1188 BP.
AC V62395;
DT 02-FEB-1999 (first entry)
DE Human Desert hedgehog protein cDNA.
KW Desert hedgehog; HHH; human; ds.
OS Homo sapiens.
FH Key location/Qualifiers
FT sig_peptide 1..66
FT mat_peptide 67..594

```

```

FT EP-824048-A2 /tag- b
PN 28-OCT-1998
PR 24-APR-1998 303187.
PR 14-APR-1997: JP-117873.
PR 25-APR-1997: JP-121578.
PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
PI Ariyasu T, Nakamura S, Orita K;
DR WP1: 98-544642/47.
DR P-PSDB: W79595.
PT Human Desert hedgehog protein - and corresponding DNA and monoclonal
PT antibody
PS Claim 10; Page 24-25; 39pp; English.
CC This cDNA sequence codes for a precursor (see W79595) of a novel
CC human Desert hedgehog protein (see also W79593). Human Desert
CC hedgehog protein DNA (HHH) was originally identified following
CC screenings of human cell lines by PCR using primers (see V62402-03)
CC based on the mouse Desert hedgehog gene. These screenings
CC indicated that the ARH-77 (ATCC CRJ-1621) cell line, which is
CC derived from the plasma cell of a leukemia patient, expressed a
CC specific gene at an elevated level. Sequencing confirmed it to
CC be a novel human gene showing homology to the mouse Desert hedgehog
CC gene. DNA encoding N-terminal sequences (see V62404-06). The invention
CC by further PCR amplifications (see V62404-06). The invention
CC provides Desert hedgehog polynucleotides and protein, a monoclonal
CC antibody (Mab) that recognises the protein, a process for producing
CC the protein, and a method for detecting the protein using the Mab.
CC The hedgehog protein, DNA and Mab can be used to elucidate
CC hereditary morphological abnormalities in humans to establish their
CC treatments and diagnoses.
SQ Sequence 1188 BP; 179 A; 376 C; 419 G; 214 T;

Query Match 98.8%; Score 1176.8; DB 1; Length 1188;
Best Local Similarity 99.4%; Pred. No. 7.7e-236;
Matches 1181; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATGGCTCTCCAGCAAACTGTGCTACAGCGCCCTGAGCACTGTGTTCGCTCGAGAGCGGCG 60
DB 1 ATGGCTCTCCAGCAAACTGTGCTACAGCGCCCTGAGCACTGTGTTCGCTCGAGAGCGGCG 60
QY 61 CAGAGCTGGGGGCGGGGCGGGGCGGGTGGCGGCGCGCTATGCGCGCAACAGACTC 120
DB 61 CAGAGCTGGGGGCGGGGCGGGGCGGGTGGCGGCGCGCTATGCGCGCAACAGACTC 120
QY 121 GTGGCGTACTCTACAGCAATTTGTGCGCGGCTGCCAGAGCGGAGCTTGGGCGCACT 180
DB 121 GTGGCGTACTCTACAGCAATTTGTGCGCGGCTGCCAGAGCGGAGCTTGGGCGCACT 180
QY 181 GGGCCAGCGAGAGGAGGTGGCAAGGGGCTCCAGGCGCTTCGGGAGCTGTGCCAAC 240
DB 181 GGGCCAGCGAGAGGAGGTGGCAAGGGGCTCCAGGCGCTTCGGGAGCTGTGCCAAC 240
QY 241 TACAACCCCGACATCTTCAAGATGAGAGAAAGTGAAGCGACCGGCTGATGACC 300
DB 241 TACAACCCCGACATCTTCAAGATGAGAGAAAGTGAAGCGACCGGCTGATGACC 300
QY 301 GAGGCTTGCAGAGAGAGGTGAAGCGCTTTGGCCATTTGCCGTGATGAACATGTGCCGGA 360
DB 301 GAGGCTTGCAGAGAGAGGTGAAGCGCTTTGGCCATTTGCCGTGATGAACATGTGCCGGA 360
QY 361 GTGGCCCTACAGAGTACTGAGAGGCTGGGACAGAGAGGCGCACAGGCTCAGATTACTC 420
DB 361 GTGGCCCTACAGAGTACTGAGAGGCTGGGACAGAGAGGCGCACAGGCTCAGATTACTC 420
QY 421 CACTACGAAGGCGGTGCTTTGGACATCACTAGTGTGACCGCGACCGCAACAGATGGG 480
DB 421 CACTACGAAGGCGGTGCTTTGGACATCACTAGTGTGACCGCGACCGCAACAGATGGG 480
QY 481 TTGCTGGCGCGCCCTTCGACAGTGAAGCCGGCTTCGACTGGGTCTACTACGATCCCGAAC 540
DB 481 TTGCTGGCGCGCCCTTCGACAGTGAAGCCGGCTTCGACTGGGTCTACTACGATCCCGAAC 540

```


Query Match	Best Local Similarity	Score	DB 1	Length
Matches 1052; Conservative 0; Mismatches 139; Indels 0; Gaps	81.3%;	968.6;	DB 1;	1191;

1 ATGGCTCTCCCTGACCAATCTACTGCTTGTGCTGCTTGGACACTTCTGGCGCTGCCAGCC 60

Dp	1	ATGGCTCTGCGGGCCAGTGTGTGGCCCTGTGCTCTGGACACTTGTGGCACTATCTGCC	60
QY	61	CAGAGCTGCGGGCGGGGCGGGGCGGTTGGCGGGCGCGCTATGCGCGCAAGCACTC	120
Dp	61	CAGAGCTGCGGGCGGGGCGGGGCGGGTGGCGGGCGGTTATGTGCGCAACACTT	120
QY	121	GTGCGGCTACTCTACACGAATTTGTGCCGGCGTGGCCAGAGCGGACCTTGGGCGCACT	180
Dp	121	GTGCGCTCTGTATACAGCAATTTGTGCCCACTATGACCGAGCGGACCTTGGGCGCACT	180
QY	181	GGGCGACGGAGGGGGGGTGGCAAGGGGCTCCGAGGCTTCCGGGACCTGTGCTGCCAAC	240
Dp	181	GGGCGACGGAGGGGGGGTGAACAGGGGCTGGAGCGCTTCCGGGACCTGTGACCCAAC	240
QY	241	TACAACCCCGACATCTCTTCAAGATGAGAGAAACAGTGGAGCGGACCGGCTGATGACC	300
Dp	241	TACAACCCCGACATATCTTCAAGATGAGAGAAACAGGGGGGACGACGCGCTGATGACA	300
QY	301	GAGCGTTGCAGAGAGAGGGTGAACCGTTTGGCCATTGCCGTGATGACATGTGGCCGGA	360
Dp	301	GAGCGTTGCAGAAAGGGGGGTGAACCGTTCTACGACATCGGGGTATGATGACATGTGGCCGGA	360
QY	361	GTGGCGCTACGATGACTGAGGGGCTGGGCGAGAGACGGCGACACAGCTCAGAGATTCACTC	420
Dp	361	GTAGCGCTACGATGTGACTGAGGGCTGGGCGAGAGACGGCGACACAGGATTCACCTC	420
QY	421	CACCTACGAAGCGCGTCTTTTGGACATCACTATGCTGTGACCGCGACCGCAACAGATTGGG	480
Dp	421	CACCTACGAAGCGCGTCTTGGACATCACTATGCTGTGACCGCGACCGATATGATGATGGT	480
QY	481	TTCGTGGCGCGCTCCGCAATGGAGCGCGGCTTCGACTGGGTCTACTACGAGTCCCGCAAC	540
Dp	481	TTCGTGGCGCGCTCTACTGTGGAAGCGGATTCGACTGGGTCTACTACGAGTCCCGCAAC	540
QY	541	CAGTCCAGCTGTGCGTCAAAAGCTGATTACTACAGGGGGTCCGGGGGGGGCGGCTT	600
Dp	541	CACCTCACGATATGCGTCAAAAGCTGATTACTACAGGGGGTCCGGGGGGGGCGGCTT	600
QY	601	CCGGGAATCAACTCTGCGGCTGTGGACCGCGCGAGCGGAAGGCGTCCGGAATCGCAC	660
Dp	601	CCGGGAATCCACAGGCGGTGGCTTGGGACCGCGCGAGCGGAAGGGGCTAGGGAATCAT	660
QY	661	CGGCGAGCTGGGTTTGGGCGCGCATGCGTGAAGCGCGGGTGTGCCACGCGGCTGCTG	720
Dp	661	CGTGTGACTGGGTACTGAGCGCGCTGATGAGAGGGGCGGAGTGTGCCACCGCACTGCTG	720
QY	721	CTCTCTCTGAGACGGGACCTTGACAGCGCGGGGCTTATTTGTGGCGTGTGAGACGAGTGG	780
Dp	721	CTCTCTCTGAGACGGGATCTGACAGCGCGCGCGCTGTGCTGTGAGACGAGCGG	780
QY	781	CTCTCCACGCAAACTGTGTGCTACAGCGGCTTGGACCTGTGCTGTGAGAGGCGCGCG	840
Dp	781	CTCTCCGCGCAAACTGTGTGCTACACCGCTTGGCATCTGTGTTCGTGCTGGGGGCGACAGG	840
QY	841	CCGCGCCGACGCACTTTGACACCGGTGTCGCGCGCGCGCTACGCGCTGCGGGACTCGGTG	900
Dp	841	CTCTCTCTCAAGTACTTTTGGACCGGCTTTCGCGCGCGCTTACGTGCTGGGCACTCGGTG	900
QY	901	CTGGCGCGCGCGGGGATGCGTGTGGGCCAAGGCGCGGTGCGCCGCTGTGGCGGGGAGGA	960
Dp	901	CTGGGCTCCCGCGGGGACGCGGCTCCAGCGCGCGCGGTGACCCGCGTGGCGGCGAGGA	960
QY	961	GCGCTGGGCGCTGTTCGCGCGCTACACCGGACGAGCTGTGTCAAGAGCTCTC	1020
Dp	961	GCGCTGGGCGCTGTTCGACCGCTCACTGGGCAAGGGAGCGTGTGTCAAGAGCTCTC	1020
QY	1021	GCTCTTGTCTACGCGGTTCTTGAGAGTCAACACAGTGGGCGACCGCGCTTTTGGCCCCCTTG	1080
Dp	1021	GCTCTCTGTCTACGCGGTTCTAGAGAGTCAACACAGTGGGCGCACCGCGCTTTCGCCCTTTTG	1080
QY	1081	AGATGCTGTACAGCGGCTAGGGGGGCTGTCTCCCGGGGGGCGGTCCACACGACTGTGATG	1140
Dp	1081	AGATGCTGTACAGCGGCTAGGGGGGCTGTCTCCCGGGGGGCGGTCCACACGACTGTGATG	1140

Db 1081 CGCCTGCTGACACGGCTGGGGCTCTGCTCCCTGGGGGTGCAATGCCACCTAGTGGCATG 1140
QY 1141 CATGTGACTCTCGGCTCTCTACCGCTTACGAGAGAGACTAGTGGGCTGA 1191
|||||
Db 1141 CATGTGACTCTCGGCTCTCTTACCGCTTGGCCGAGAGATTAATGGGCTGA 1191
RESULT 6
QY1642
ID 091642 standard; cDNA; 1190 BP.
AC 091642;
DT 18-MAR-1996 (first entry)
DE Mouse desert hedgehog protein gene.
KW Mouse; desert hedgehog gene; probe; primer; diagnostic;
KW nervous system disorder; gene therapy; antibody; ds.
OS Mus musculus.
FH Key 1. .303 Location/Qualifiers
FT exon 304. .567 /*tag= a
FT 568. .1188 /*tag= b
FT exon 568. .1188 /*tag= c
PN MO9518856-A1.
PD 13-JUL-1995.
PF 30-DEC-1994; U14992.
PR 30-DEC-1993; US-176427.
PR 14-DEC-1994; US-356060.
PA (HARD) HARVARD COLLEGE.
PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
PI Ingham PW, McMahon AP, Tablin CJ;
DR WPI: 95-255060/33.
DR P-PSDB: R77345.
PT Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful
PT to treat degenerative nervous system disorder(s) and in gene
PT therapy.
PS Claim 5; Page 135-37; 210pp; English.
CC The sequence encodes a mouse desert hedgehog protein, homologous
CC to a Drosophila hedgehog protein (R77337), and has been isolated by
CC low stringency screening of a mouse genome DNA library, in phage
CC lambda. The sequence contains 3 homologous regions, encoding a
CC single open reading frame interrupted by introns. Splicing has
CC been confirmed by polymerase chain reaction amplification of first
CC strand cDNA generated from adult testicle RNA. Probes and primers
CC derived from hedgehog sequences may be used as diagnostic agents
CC for neuromuscular, autonomic or central nervous system disorders,
CC and the gene may also be used in gene therapy. Antibodies
CC generated from the encoded protein may be used as therapeutic or
CC research reagents.
SQ Sequence 1190 BP; 194 A; 371 C; 399 G; 226 T;
Query Match 81.2%; Score 967.6; DB 1; Length 1190;
Best Local Similarity 88.3%; Pred. No. 1.8e-192;
Matches 1051; Conservative 0; Mismatches 139; Indels 0; Gaps 0;
QY 1 ATGGCTCTCTGACCAATCTACTGCGCTTGTGCTGGCACTTGGCGCTGCCACC 60
|||||
Db 1 ATGGCTCTGCGGCGACAGTGTGCTGCTGTGCTGGCACTTGGCACTTATCTGCC 60
QY 61 CAGAGCTGCGGGCGGGCGGGCGGGTGTGGCGGGCGCGCTATGCGCGAACAAGCTC 120
|||||
Db 61 CAGAGCTGCGGGCGGGCGGGCGGGCGGGTGTGGCGGGCGCGCTTATGTGCGCAACAAGCTT 120
QY 121 GTGCGGCTACTCTTACAAAGCAATTTTGGCCGGCGGCGCAAGACCGCTTGGGGCCAGT 180
|||||
Db 121 GTGCGCTCTGCTATACAAAGCAATTTTGGCCAGATATGCCGAGGACCTTGGGGCCAGT 180
QY 181 GGGCGAGCGAGGAGGAGGTGCGAAGGGCTCCGAGCGCTTCCGGGAACTGTGCCCAAC 240
|||||
Db 181 GGGCGAGCGAGGAGGAGGTGCGAAGGGCTCCGAGCGCTTCCGGGAACTGTGCCCAAC 240
QY 241 TACAACCCCGACATCATCTTCAAGAGATGAGAGAACTGAGGCCGCGCTGATGACC 300

|||||
Db 241 TACAACCCCGACATCAATCTTCAAGAGATGAGAGAAAGCGCCGACAGCCCTGATGACA 300
QY 301 GAGCGTTGCAAGAGAGAGGTGAACGCTTTGGCCATTGCTCGTATGAACATGTGGCCGGA 360
|||||
Db 301 GAGCGTTGCAAGAGAGCGGGTGAACGCTTGAACCATCGCGGTATGAACATGTGGCCGGA 360
QY 361 GTGCGCTTACGAGTACTGAGGGCTGGGAGAGAGACGGCCAGCAAGCTTCAAGATTCACTC 420
|||||
Db 361 GTAGCGCTTACGAGTACTGAGGGCTGGGAGAGAGAGCGCCAGCACAGCAAGATTCACTC 420
QY 421 CACTACGAAGCGCGTGTGTTGGACATCACTACGTCTTACCGCGGCAAGCAAGTATGGG 480
|||||
Db 421 CACTACGAAGCGCGTGTGTTGGACATCACTACGTCTTACCGCGTATTAATGATAGT 480
QY 481 TTGCTGCGCGCCCTCCGAGTGAAGCCGGCTTGGACTGAGTCTACTACGAGTCCCGAAC 540
|||||
Db 481 TTGTTGGCGCGCCCTAGCTGATGAAGCCGGATTGACTGGGTCTACTACGATCCCGAAC 540
QY 541 CACGTCCAGCTGTGCTCAAAAGCTGATTACTCACTGAGGCTCCGGGCGGGCGCTGCTT 600
|||||
Db 541 CACATCCAGCTATCGGTCAAAAGCTGATTACTCACTGAGGCTCCGAGCGGAGGCTGCTT 600
QY 601 CCGGGAATGCAACTGTGCGCCCTGTGAGCGCGGAGCGGAGAAAGGCTGCGGAACTGAC 660
|||||
Db 601 CCGGGAATGCAACTGTGCGCCCTGTGAGCGCGGAGCGGAGAAAGGCTGAGGAACTGAC 660
QY 661 CGCGAGACTGGGTTTGGGGCGGATGCTGAGCGCGGGGTGGTCCGACGCGGTGCTG 720
|||||
Db 661 CGTGTGTAAGTGGTACTGCGCGCTGATGCAAGCGGCGGAGTGGTACCACCGCACTGCTG 720
QY 721 CTCTCTGGAACGGGACTTGAAGCGCGGCGCTTCAATTTGTGCTGTGAGAACGATGG 780
|||||
Db 721 CTCTCTGGAACGGGACTTGAAGCGCGGCGCTTCAATTTGTGCTGTGAGAACGATGG 780
QY 781 CCTCCAGGCAACTGTGCTGACAGCCCTGTGCACTGCTGTTCCCTGAGAGCGCGCG 840
|||||
Db 781 CCTCCAGGCAACTGTGCTGACAGCCCTGTGCACTGCTGTTCCCTGAGAGCGCGCG 840
QY 841 CCGCGGCGGAGGCACTTGTGACCGGTTGTGCGCGCGCGGCTGACGCGTGGGAGTGGT 900
|||||
Db 841 CCGCGGCGGAGGCACTTGTGACCGGTTGTGCGCGCGCGGCTGACGCGTGGGAGTGGT 900
QY 901 CTGGCGCCCGGGGGGATGCGCTTGGCGCAAGCGCGCGTGGCGCGTGTGGCGGGAGAA 960
|||||
Db 901 CTGGCGCCCGGGGGGATGCGCTTGGCGCAAGCGCGCGTGGCGCGTGTGGCGGGAGAA 960
QY 961 GCGTGGGCGGTTTGGCGCGGCTACCGCGGACGAGGAGAGTGTGGAGACATGTCCTG 1020
|||||
Db 961 GCGTGGGCGGTTTGGCGCGGCTACCGCGGAGCGGAGAGTGTGGAGACATGTCCTG 1020
QY 1021 GCGTCTGTACGCGGTTGTGAGAGATCACCAAGTGGGCGGCGGTTTGGCCCTTGG 1080
|||||
Db 1021 GCGTCTGTACGCGGTTGTGAGAGATCACCAAGTGGGCGGCGGCTTGGCCCTTGG 1080
QY 1081 AGACTGTGACAGCGGCTAGGAGGCGTGTCCCGCGCGGGCGCTCCAGCGGACTGGCATG 1140
|||||
Db 1081 CCGCTGTGACAGCGGCTAGGAGGCGTGTCCCGCGGGGTGCAATCCAGCGGACTGGCATG 1140
QY 1141 CATTTGACTCTGCGCTCTCTACCGCTTACGCGGAGAGAGTACTGGGCTG 1190
|||||
Db 1141 CATTTGACTCTGCGCTCTCTTACCGGTTTGGCGAGAGATTAATGGGCTG 1190
RESULT 7
X16183
ID X16183 standard; cDNA; 1190 BP.
AC X16183;
DT 29-APR-1999 (first entry)
DE Mouse dbh hedgehog cDNA sequence.
KW Patched; hedgehog; ptc therapeutic; neuroprotective; neuronal cell;
KW brain infarction; cerebral infarction; transient ischemic attack;
KW stroke; cerebral infarct volume; spinal cord; oedema; trauma;

PT Inhibitor, or an agent derived from hedgehog polypeptides, useful in
PT the treatment of Parkinson's disease
PS Disclosure; Page 85-86; 138pp; English.
CC This nucleotide sequence comprises a coding region for the mouse
CC Dhh Desert hedgehog protein (see W97766). The invention is based on
CC the finding that hedgehog proteins are useful as protective agents
CC in the treatment and prophylaxis of neurodegenerative disorders
CC resulting from the loss of dopaminergic and/or GABA-ergic neurons,
CC or the general loss of tissue from the substantia nigra.
CC Exemplary disorders include Parkinson's disease, Huntington's
CC disease (both claimed), amyotrophic lateral sclerosis and cerebral
CC ischaemia. The invention relates to hedgehog therapeutics (i.e.
CC hedgehog polypeptides and gene therapy constructs e.g. constructs
CC encoding recombinant hedgehog polypeptides and trans-activation
CC constructs for altering hedgehog gene regulatory sequences) and
CC ptc therapeutics (i.e. agents which mimic the effect of naturally
CC occurring hedgehog proteins on patched signalling) that are
CC effective in both human and animal subjects. Human Ihh and Dhh
CC polypeptides (see W97763-64) are preferred. The products can also
CC be used for the maintenance of differentiated neurons in cultures,
CC and to enhance the implantation of such neuronal cells in an
CC animal. They can also be used to prevent or treat neurodegenerative
CC conditions arising from the use of certain drugs, and in the
CC prevention and/or treatment of hypoxia, e.g. as a neuroprotective
CC agent.
Q Sequence 1190 BP; 194 A; 371 C; 399 G; 226 T;

Query Match	81.2%;	Score 967.6;	DB 1;	Length 1190;
Best Local Similarity	88.3%;	Pred. No. 1.8e-192;		
Matches 1051;	Conservative 0;	Mismatches 139;	Indels 0;	Gaps 0;

OY	1	ATGGCTCCCTGACCAATCTACTGCGCCCTTGTGCGCTTGAGCACTTTCGGGCTCCGAGGC	60
Db	1	ATGGCTCGCGGGCCGACGCTCTGTGGCCCTGTGCTGCTTGGCACTTGTGGCACTATCTGGC	60
OY	61	CAGAGCTGCGGGCCCGGCGCGGGCCGGTGTGGCCCGGCCCTTATGCGCGCAACCACTC	120
Db	61	CAGAGCTGCGGGCCCGGCGCGAGAACCGGTTGTGGCGGGCGGCTTATGTGGCGCAACCACTT	120
OY	121	GTCGCCGCTACTCTACAGCAATTTGTGCCCCGGCGCTGCCAGAGCGGACCCCTTGGCGCCAGT	180
Db	121	GTCGCTCTGCTATACAAACCACTTTGTGCCAGTATGCCCCGAGCGGACCTTGGGCGCAGT	180
OY	181	GGGCGCAGCGGAGGGGAGGGGTGGCAAGGGGGCTCCGAGCGCTCCGCGAGCCTGTGTCGCAAC	240
Db	181	GGCGCAGCGGAGGGGAGGGGTAAACAAGGGGGGTGCGAGCGCTTCCGGGACTGTATCCCAAC	240
OY	241	TACAACCCCGCATCATCTTCAAGGATAGAGAGAACAAGTGGAGCGCACCGCGCTGATGACC	300
Db	241	TACAACCCCGCATATATCTTCAAGGATAGAGAGAACAACGCGCGACAGCCGCTGATGACA	300
OY	301	GAGGTTTCCAGGAGAGGGGTGAAGGCTTGTGGCCATTTGCCGTGATGAAATATGTGGCCCCGA	360
Db	301	GAGGTTTCCAAAGAGCGGGGTGAAGGCTTGTACCAATCGGGGTGAATGAACATTTGGCCCCGA	360
OY	361	GTGGCGCTACAGATGACTGAGGGCTTGGGACAGAGACGGCCACACGCTCAGAGATTCACTC	420
Db	361	GTAGCGCTTACTGTGACTGGAAGGCTGGACAGAGACGGCCACACGCAAGATTCATCTC	420
OY	421	CACCTACGAAGGCGGTGCTTTSGGACATCACTAGTCTGACCGCGACCGCACMACMAATATGGG	480
Db	421	CACCTACGAAGGCGGTGCTTGTGACATCACTAGTCTGACCGGTGACCGGTAAATAATATGGT	480
OY	481	TTTGCTGGGCGCCCTTCGCACTGTAAGACCGGGCTTTCGACTGTGCTACTACAGATCCCCGCAAC	540
Db	481	TTTGTTGGGCGCCCTTACGCTGTGAAGCCGGATTTCAGCTGGGCTCTACTACGAATCCCCGCAAC	540
OY	541	CACGTCACAGTGTGCGGTAAAGCTGATTAACCACTGAGGGGGTCCGAGGGGGGGGGGTGCTTT	600
Db	541	CACATCCACGATGTGGGTCAAAAGCTGATTAACCACTGGCGGTCCAGCGCGGAGGGCTGCTTT	600
OY	601	CCGGGAATTCCACTGTGCGCTGTGAGCGCGCAGCGGAAGGGCTTGCAGGAATCTCAC	660

Db	601	CCGGGAATGCAACGCGTCCCTTGGAGCGCGCAAGGAGAGGGCTTAGGGAACTACT	660
Qy	661	CACGGAAGCTGGTTTTTGGCGGCCGATGCGTACAGCCGGGTGTGCCAGCGCGTCTG	720
Db	661	CGTGGTGCATGGGTAACGAGCCGCTGATGACAGCGGCCGAGTGTATCCACGCCAATGCTG	720
Qy	721	CTCTTCCGAGACCGGGACTTGCAGCGCCGGGGCTCAATTGTGGCTGTGAGACCGAATGG	780
Db	721	CTCTTCCTGGAGCGGGATCTGCAAGCGCCGGGCTCGTTGTGGCTGTGAGACCGAGCGG	780
Qy	781	CCTCCACGCAAACTGTGTGCTACAGCCCTGAGCACTTGTGTTGCCGCTGAGAGCGCCGCG	840
Db	781	CCCTCCGGCAAACTGTGTGCTCACACCCCTGGCATCTGTGTTCGTGCTCGGGGGCAAGC	840
Qy	841	CCCCGGCCAGCACTTTGCACCGGTTGTCGGCGCCGCGCTACGGCTTGGGAGACTCGGTG	900
Db	841	CCCTCTCCAGGTGACTTTGCACCGGGTTTCGGCGCCGCTTACGTGCTGGGAACTCGGTG	900
Qy	901	CTGCGCCCGCGGGGAGTGGGCTTCCGCCAGCGAGCGGGGTGGCCGTGAGCGGGAGAGA	960
Db	901	CTGGCTCCCGCGGGAGACGGGCTTCAGCCGGCGCGGTATCCCGGGTGGCCGCGAGAGA	960
Qy	961	GCCGTGGGCGTGTCCGCGCGCTACCCGCGCAAGGGAGCGTGTGTGTAAGATGTCTGTG	1020
Db	961	GCCGTGGGCGGTGTCCGACCGCTACACTCGCGCAAGGAGACCTCTGTATCAAGACTCTTC	1020
Qy	1021	GCCCTTGTACTAGCGGGTTCGTGGAGACTCAACAGTGGGGCCACGGGGCTTTGGCCCCCTG	1080
Db	1021	GCCCTCTGTACTAGCGGGTTCAGAGACTCAACAGTGGGGCCACCGGGCTTGGCCCCCTTG	1080
Qy	1081	AGACTGCTGCACGGCGCTAGAGGGCGCTGCTCCCGGGGGGCGCGTACAGCCGACTGGCAT	1140
Db	1081	CGGCTGGCGACCGGGCGTCCGGGGCTGTGCTCCCGGGGGGTGCAAGTCCAGCCGACTGGCAT	1140
Qy	1141	CATTGTGACTCTGGGCTCTCTTACCGCTTACGGAGAGGACTTCTGGGCTG	1190
Db	1141	CATTGTGACTCTGGGCTCTTACCGCTTACCGCTTACGGAGATATATGGGCTG	1190

RESULT 9

ID V62397 standard; cDNA; 602 BP.

DT 02-FEB-1999 (first entry)

Desert hedgehog; HUDHN; human; ds.

Key	Location/Qualifiers
00	name of person:
FH	

```

      /tag= a
      FT
      FT

```

```

      / *tag= b
FT

```

PD 28-OCT-1998.

PR 14-APR-1998; JP-117873.

PA (HAYB) HAYASHIBARA SEIBU

DR WPI; 98-544642/47.

PT Human Desert hedgehog pro

PS Example 1-2; Page 27-28;

CC a precursor form (see W79

cc originally identified in

CC V62405-06) yielded cDNA e

CC W79593-95) of human Deser

CC antibody (Mab) that recognises the protein, a process for producing
 CC the protein, and a method for detecting the protein using the Mab.
 CC The hedgehog protein, DNA and Mab can be used to elucidate
 CC hereditary morphological abnormalities in humans to establish their
 CC treatments and diagnoses.
 CC Sequence 602 BP; 111 A; 187 C; 199 G; 105 T;

Query Match 49.2%; Score 586.4; DB 1; Length 602;
 Best local Similarity 99.0%; Pred. No. 1.7e-113;

Matches 590; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ATGGCTCTCTGACCAATCTACTGCCCCCTGTGCTGCTTGGCACTTGTGGCGTCCAGCC 60
 Db 7 ATGGCTCTCTGACCAATCTACTGCCCCCTGTGCTGCTTGGCACTTGTGGCGTCCAGCC 66
 QY 61 CAGAGTGG 120
 Db 67 CAGAGTGG 126
 QY 121 GTGCCCTACTCTACAAGCAATTTGTGCCGGGTGCGCAGAGGCACTTGGGGCGCAGT 180
 Db 127 GTGCCCTACTCTACAAGCAATTTGTGCCGGGTGCGCAGAGGCACTTGGGGCGCAGT 186
 QY 181 GGGCCAGCGAGGGGAGGGTGGCAAGGGGCTCCAGCGCTTCCGGGACCTGTGCCAAC 240
 Db 187 GGGCCAGCGAGGGGAGGGTGGCAAGGGGCTCCAGCGCTTCCGGGACCTGTGCCAAC 246
 QY 241 TAAACCCCGACATCTATCTTCAAGATGAGAGAAACAGTGGAGCCGACCCCTGATGAC 300
 Db 247 TAAACCCCGACATCTATCTTCAAGATGAGAGAAACAGTGGAGCCGACCCCTGATGAC 306
 QY 301 GAGCGTTGCAAGAGAGGGGGAACGCTTGGCCATTGCCGTGATGAACATGTGGCCGGA 360
 Db 307 GAGCGTTGCAAGAGAGGGGGAACGCTTGGCCATTGCCGTGATGAACATGTGGCCGGA 366
 QY 361 GTGCGCTTACGATGACTGAGGGCTGGAGAGAGAGCCGACCAAGCTCAGATTCACTC 420
 Db 367 GTGCGCTTACGATGACTGAGGGCTGGAGAGAGAGCCGACCAAGCTCAGATTCACTC 426
 QY 421 CATACGAAAGGGCGTGTGGTACATCTAGCTTGAACCGGACCGCAACAAATATGG 480
 Db 427 CATACGAAAGGGCGTGTGGTACATCTAGCTTGAACCGGACCGCAACAAATATGG 486
 QY 481 TTGCTGGCGGCGCTTCGAGTGAAGCGGCTTGGACTGGGTCTACTACGAGTCCCGAAC 540
 Db 487 TTGCTGGCGGCGCTTCGAGTGAAGCGGCTTGGACTGGGTCTACTACGAGTCCCGAAC 546
 QY 541 CACGTCACAGTGTGCGTCAAAAGCTGATTACTACTGGGGTCCGGGGGGGGCTG 596
 Db 547 CACATCCACAGTGTGCGTCAAAAGCTGATTACTACTGGGGTCCGGGGGGGGCTG 602

RESULT 10

V62398
 ID V62398 standard; cDNA; 575 BP.

AC V62398;
 DT 02-FEB-1999 (first entry)
 DE Human Desert hedgehog protein C-terminal region cDNA.
 KW Desert hedgehog; HuhDH; human; ds.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT mat_peptide 2..574
 PN /*tag- a
 PR 24-OCT-1998.
 PR 24-APR-1998; 303187.
 PR 14-APR-1998; JP-117873.
 PR 25-APR-1997; JP-121578.
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PI Ariyasu T, Nakamura S, Orita K;
 DR MPI; 98-544642/47.
 P-PSDB; W79598.

PT Human Desert hedgehog protein - and corresponding DNA and monoclonal
 PT antibody
 PS Example 1-3: Page 28-29; 39pp; English.

CC This cDNA sequence codes for the C-terminal region (see W79598) of
 CC a precursor form (see W79594) of novel human Desert hedgehog protein
 CC (see also W79593). Human Desert hedgehog protein DNA (HuhDH) was
 CC originally identified in human following ARH-77 (ATCC CRL-1621) cell
 CC line following RT-PCR amplification screenings. Further PCR (see
 CC W62407-08) yielded cDNA encoding the C-terminal region. DNA
 CC molecules (see W62393-95) encoding mature and precursor forms (see
 CC W79593-95) of human Desert hedgehog are claimed, as are a monoclonal
 CC antibody (Mab) that recognises the protein, a process for producing
 CC the protein, and a method for detecting the protein using the Mab.
 CC The hedgehog protein, DNA and Mab can be used to elucidate
 CC hereditary morphological abnormalities in humans to establish their
 CC treatments and diagnoses.
 CC Sequence 575 BP; 66 A; 183 C; 221 G; 105 T;

Query Match 48.1%; Score 573.4; DB 1; Length 575;
 Best local Similarity 99.8%; Pred. No. 8.2e-111;

Matches 574; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 549 CGTGTGGTCAAGCTGATTACTACTGGCGGTCCGGGGGGGGCTGCTTCCGGAAA 608
 Db 1 CGTGTGGTCAAGCTGATTACTACTGGCGGTCCGGGGGGGGCTGCTTCCGGAAA 60
 QY 609 TGCATGTGCGGCTGTGAGCGGGGCAAGGGAAGGGCTCCGGGAATGACCGCGGAGA 668
 Db 61 TGCATGTGCGGCTGTGAGCGGGGCAAGGGAAGGGCTCCGGGAATGACCGCGGAGA 120
 QY 669 CTGGGTTTGGCGCGGATGCGTCAAGCGGGGTGGTGGCCACGCGGTGCTTCTCT 728
 Db 121 CTGGGTTTGGCGCGGATGCGTCAAGCGGGGTGGTGGCCACGCGGTGCTTCTCT 180
 QY 729 GAGCGGGGACTTGCAGCGCGGGCTTCATTGTGCTGTGAGACCGAGTGGCTCCAG 788
 Db 181 GAGCGGGGACTTGCAGCGCGGGCTTCATTGTGCTGTGAGACCGAGTGGCTCCAG 240
 QY 789 CAACATGTTGCTACGCGCCCTGGGACCTGTGTTGGCGCTCAGAGGGCCGGCGCC 848
 Db 241 CAACATGTTGCTACGCGCCCTGGGACCTGTGTTGGCGCTCAGAGGGCCGGCGCC 300
 QY 849 AGCGCATTTGACCGCGATGTTCGCGCGCGGCTACGCGGCTGGGACTCGGTGCGCC 908
 Db 301 AGCGCATTTGACCGCGATGTTCGCGCGCGGCTACGCGGCTGGGACTCGGTGCGCC 360
 QY 909 CGCGGGGATGCGCTTCGCGCAGCGCGGTGCGCCGTGTGGCGCGGAGAGACCGTGG 968
 Db 361 CGCGGGGATGCGCTTCGCGCAGCGCGGTGCGCCGTGTGGCGCGGAGAGACCGTGG 420
 QY 969 CGTGTTCGCGCGCTCAACCGCGGACGCTGTGTGTAACGATGCTTGGCTCTTG 1028
 Db 421 CGTGTTCGCGCGCTCAACCGCGGACGCTGTGTGTAACGATGCTTGGCTCTTG 480
 QY 1029 CTACGGGGTCTGAGAGTCAACAGTGGGGGACCGCGCTTTGGCCCTTGAAGTGT 1088
 Db 481 CTACGGGGTCTGAGAGTCAACAGTGGGGGACCGCGCTTTGGCCCTTGAAGTGT 540
 QY 1089 GCACGGGCTAGAGGGCGCTGCTCCCGCGGGGGCGG 1123
 Db 541 GCACGGGCTAGAGGGCGCTGCTCCCGCGGGGGCGG 575

RESULT 11

V62396
 ID V62396 standard; cDNA; 548 BP.

AC V62396;
 DT 02-FEB-1999 (first entry)
 DE Human Desert hedgehog protein cDNA clone pHUHH/#20.
 KW Desert hedgehog; HuhDH; human; ds.
 OS Homo sapiens.
 FH Key Location/Qualifiers

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2000, 08:17:52 ; Search time 22.84 Seconds
(without alignments)
1016.523 Million cell updates/sec

Title: US-08-900-220-17

Perfect score: 2088
Sequence: 1 MALNTNLPCLCCALLALP...PTGMHYSRLYRLAEELG 396

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 168808 segs, 58629743 residues

Total number of hits satisfying chosen parameters: 168808

Minimum DB seq length: 0
Maximum DB seq length: 100000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : PIR_63:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2008	96.2	396	B49425	Desert hedgehog pr
2	1217	58.3	437	A49425	Sonic hedgehog pro
3	1203	57.6	437	B53193	hedgehog homolog v
4	1198	57.4	425	A49424	patterning protein
5	1120	53.6	418	A53193	hedgehog homolog v
6	1092.5	52.3	444	S56765	morphogen Xhh prec
7	1069	51.2	415	A49426	sonic hedgehog gen
8	956.5	45.8	336	C49425	Indian hedgehog pr
9	951	45.5	471	A46400	segment polarity p
10	495	23.7	94	G02735	desert hedgehog -
11	209	10.0	868	T22281	hypothetical prote
12	194.5	9.3	1226	T24045	hypothetical prote
13	193.5	9.3	1021	T23252	hypothetical prote
14	186.5	8.9	1207	T23754	hypothetical prote
15	182	8.7	481	T27665	hypothetical prote
16	163	7.8	484	T34504	hypothetical prote
17	161.5	7.7	629	T19563	hypothetical prote
18	150.5	7.2	615	T29550	hypothetical prote
19	135	6.5	205	T26250	hypothetical prote
20	117	5.6	3172	S22012	erythronolide synt
21	115.5	5.5	3670	T36242	cap peptide synthe
22	114.5	5.5	687	T08528	probable DNA topoi
23	112.5	5.4	846	Q08EC3	HQRF1 protein - hu
24	111.5	5.3	788	Q08EC3	HQRF1 protein - hu
25	111.5	5.3	3178	S13595	6-deoxyerythronoli
26	106.5	5.1	481	T27975	hypothetical prote
27	103.5	5.0	3391	J50219	polypotein - deng
28	103	4.9	672	T36083	hypothetical prote
29	103	4.9	3519	S43048	polyketide synthas
30	102.5	4.9	3388	GNWVP	genome polypotein

31	100.5	4.8	3391	1	GNWV26	genome polypotein
32	98	4.7	474	2	C75625	hypothetical prote
33	97	4.6	3391	1	GNWV16	genome polypotein
34	96.5	4.6	541	2	T35143	probable monooxyge
35	96	4.6	759	2	S32875	hypf protein - Rhl
36	96	4.6	3396	1	A42551	genome polypotein
37	95.5	4.6	488	2	D70614	hypothetical prote
38	95.5	4.6	1039	2	A34269	integrin alpha-2b
39	95.5	4.6	1958	2	B40505	hypothetical prote
40	94.5	4.5	376	2	F70576	hypothetical prote
41	94.5	4.5	1763	3	T17465	rifamycin polyketi
42	94	4.5	401	2	B75562	molybdopterin bios
43	94	4.5	811	2	S08579	hypothetical prote
44	93	4.5	1733	1	B45344	probable nuclear a
45	92	4.4	987	2	A75496	conserved hypothet

ALIGNMENTS

RESULT 1
B49425
Desert hedgehog protein precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Dec-1999
C:Accession: B49425
R:Echeillard, Y.; Epstein, D.J.; St-Jacques, B.; Shen, L.; Mohler, J.; McMahon, T.A.; M
Cell 75, 1417-1430, 1993
A:Title: Sonic hedgehog, a member of a family of putative signaling molecules, is imp
A:Reference number: A49425; M01D:94094334
A:Accession: B49425
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-396 <ECHO>
A:Cross-references: GB:X76292; NID:g443941; PIDN:CA53924.1; PID:g443942
C:Genetics:
A:Gene: Dhh
C:Superfamily: sonic hedgehog protein

Query Match	Score	DB 2:	Length	396:
Best Local Similarity	96.2%	Pred. No. 7.4e-153;		
Matches 382; Conservative	6;	Mismatches 8;	Indels 0;	Gaps 0;
QY 1	MALNTNLPCLCCALLALP	ASQSGPGRGRRYARKOLVPLLYKQFVCPERTGAS	60	
DB 1	MALPASLLPLCCALLALP	ASQSGPGRGRRYARKOLVPLLYKQFVCPERTGAS	60	
QY 61	GPAGRYAAGSERPDLVP	NPDIIFKDEENSGADRLMTERCKERYNALAIAYNMMPG	120	
DB 61	GPAGRYAAGSERPDLVP	NPDIIFKDEENSGADRLMTERCKERYNALAIAYNMMPG	120	
QY 121	VRLVTEGWDDEGHADDS	LYEGRALDITTSDDRRKYGGLAALNAYAGDWWYYSRN	180	
DB 121	VRLVTEGWDDEGHADDS	LYEGRALDITTSDDRRKYGGLAALNAYAGDWWYYSRN	180	
QY 181	HVHVSAYADNSLAVRAG	CGPFGNATVRLMSGERGRLRELHGDVWLAAASGRVPTPV	240	
DB 181	HIHVSAYADNSLAVRAG	CGPFGNATVRLMSGERGRLRELHGDVWLAAASGRVPTPV	240	
QY 241	LFELRDLORRASFAVET	EMPRLTLTPMHLVFAARGPAPADGFAVFARLRAGDSV	300	
DB 241	LFELRDLORRASFAVET	EMPRLTLTPMHLVFAARGPAPADGFAVFARLRAGDSV	300	
QY 301	LAPGDALRPARVARVA	REEVGFAPLTAGTLLVNDVLAASCTAVLESQWARRAPAPL	360	
DB 301	LAPGDALRPARVARVA	REEVGFAPLTAGTLLVNDVLAASCTAVLESQWARRAPAPL	360	
QY 361	RLHALGALPGAGVOP	TGMHYSRLYRLAEELG 396		
DB 361	RLHALGALPGAGVOP	TGMHYSRLYRLAEELG 396		

RESULT 2
A49425

A49425

Sonic hedgehog protein precursor - mouse

C:Species: Mus musculus (house mouse)

```
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Dec-1995
```

C:\Accession: A49425

R. Echelard, Y., Epstein, D.J.; St-Jacques, B.; Shen, L.; Mohler, J.; McManon, J.A.; McManus, J.B. Cell 75, 1417-1430, 1993

Cell 75, 1417-1430, 19

A; Title: Sonic hedgehog, a member of a family of putative signaling molecules,

A; Reference number: A49425; MUID:94094334

A;Accession: A49425

A; Status: preliminary; nucleic acid sequence not shown

A: Molecule type: mRNA

A;RESIDUES: 1-43 / <ELCH>

A/Notes: authors travel on 1
A/Cross-References: GB:X/6290

A/NOTE: AUCIORS TRANSDATION IS SHOWN FOR THE COUON FOR AT POSITION 430

A:Gene: Shh

C:\Superfam1

[illegible]

Query Match	58.3%;	Score 1217;	DB 2;	Length 437;
Best Local Similarity	58.7%;	Pred. No. 1.1e-89;		
Matches 249;	Conservative 53;	Mismatches 82;	Indels 40;	Gaps 12

[illegible]

RESULT 3
p53193
hedgenog homolog vhh-1 - rat
C:Species: Rattus norvegicus (Norway rat)
C:date: 06-Jan-1995 #sequence
C:Accession: B53193
R:Roelink, H.; Augburger, A.; Heemskerk, J.; Kozh, V.; Norlin, S.; Ruiz i Altaba, A.;
Cell 76, 761-775, 1994
A>Title: Floor plate and motor neuron induction by vhh-1, a vertebrate homolog of hedgehog
A:Reference number: A53193; MUID:94170375
A:Accession: B53193
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-437 <ROD>
A:Cross-references: G6:1D27340; NID:G452122; PIDN:AAA20999.1; PID:G452123

C;Superfamily: sonic hedgehog protein

Query Match	57.6%	Score 1203;	DB 2;	Length 437;
Best Local Similarity	57.8%	Pred. NO. 1.5e-88;		
Matches 245;	Conservative 52;	Mismatches 87;	Indels 40;	Gaps 10;

[illegible]

```

RESULT      4
A49424
    patterning protein Sonic hedgehog precursor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Dec-1995
C:Accession: A49424
R:Riddle, R.D.; Johnson, R.L.; Laufer, E.; Tablin, C.
    Cell 75, 1401-1416, 1993
A:Title: Sonic hedgehog mediates the polarizing activity of the ZPA.
A:Reference number: A49424; M0ID:94094353
A:Accession: A49424
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-425 <RTD>
A:Cross-references: GB:128099; NID:9453526; PTD:9455527
C:Superfamily: sonic hedgehog protein
F1-26/Domain: signal sequence #status predicted <SIG>

```

	Query Match	57.4%	Score 1198:	DB 2,	Length 425;
	Best Local Similarity	59.4%,	Pred.	No. 3.6e-88;	
	Matches 244; Conservative	56;	Mismatches 91;	Indels 20;	Gaps 10;
QY	1 MALLNLPL---LCCALALAPAOCSGPGRGVGRRRYARKOVLPIKYQFPGVPERTL	57			
	: : : : : :				
Db	4 MLILRILLVGRIC--ALLVSGLFCGPGRG-IKRRRHPRK-LTFLAYKQFIIPNAEKTLL	59			
QY	58 GASGPEGVAVANGSERFKDLPVNPNPDIIFKDEENSGADRLMTERCKEVNALALAVMM	117			
	: : : : : :				
Db	60 GASGRYEKITENSERFKELTNPYNPDITFKDEENTGADRMTQCKDKKLALAIISVMNQ	119			
QY	118 WEGVRILTVEGWDEDGHHAODSLHYEGRALDITTSDDRDNNKYGLLARLAVERGFDMYYE	177			
	: : : : : :				

Best Local Similarity 54.5%; Pred. NO. 6.9e-78;
Matches 211; Conservative 63; Mismatches 109; Indels 4; Gaps 4.

[illegible][illegible]

```
QY      370 LPGA VOPTGMHWYRLLYRLAEELL 395
          | :| |:| ||| ||
Db      301 TP S-----EGVHSYPQLMYRLGLLL 321
```

C:Keywords: transmembrane protein
F:62-82/Domain: transmembrane #status predicted <1MM>

```
Query Match          45.5%  Score 951:  DB 2:  Length 471:
Best Local Similarity 49.4%  Pred. No. 2,1e-68:
Matches 203:  Conservative 60:  Mismatches 122:  Indels 26:  Gaps 7:

QY      4  LTNTLPLOCIALLLAL-----PAQSCGPGRGVGRRRYARKOLVLYLKQEPGVPERTLG 58
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db       61  LSRLLSLVALLIYLPWFSPAHSCGPGRG--LGNHR--ARLLYLVLYLKQITPNNSEYTN 117

QY      59  ASGPAEGHVANGSERFNDLPVNNPDIIFKDENSGADRLTERCKERYN#LAIAYNMW 118
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db       118  ASGPLEGVIRRDSKPKFOLVNNRDIIFRDEESTGADRLMSKCKEKLNVLAISVNMW 177

QY      119  PGVRLRTVEGDEGHHARQDSLHTEGRALDITTSYDQRNNTGGLARLAVAGEFVWYTES 178
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db       178  PGIRLLVTESWDEDEYHHGQESLHTEGRAVITATSDRQSKGMLARLAVAGEFVWYSYS 237

QY      179  RNHHVSVKADNSLAVAGGCFEPMNAIVRLMSGCRKGLREIHRQDWLADASGRVYPTP 238
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db       238  RRHHYCVSKSSQSSISLHVHGCFTEPESTALLSEYGRKGLGEISTDRVLSMTANQAYSE 297
```

[illegible]

RESULT 10
G02735
desert hedgehog - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_rev1sion 06-Jun-1997 #text_change 17-Jul-1998
C:Accession: G02735
R:Drummond, T.A.
Submitted to the EMBL Data Library, June 1996
A:Reference number: H01643
A:Accession: G02735
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-94 <DRU>
A:Cross-references: EMBL:U59748; NID:g1401271; PID:g1401272
C:Genetics:
A:Gene: hdmH

	Query Match	23.7%	Score 495	DB 2	Length 94	
	Best Local Similarity	98.9%	Pred. No. 7.3e-33			
	Matches	93	Conservative	0	Mismatches 1	Indels 0
QY	85	IFKDENSGALRLTERCKERVNALAIIVMMWMPGRLRVTEGMDEDGHHADDSLHYEG	144			
Db	1	IFKDENSGADRLTERCKERVNALAIIVMMWMPGRLRVTEGMDEDGHHADDSLHYEG	60			
QY	145	RALDITTSDDRDNRKYGILARLAVEAGFDWYVES	178			
Db	61	RALDITTSDDRDNRKYGILARLAVEAGFDWYVGS	94			

```

RESULT 11
T22281
hypothetical protein F46B3.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T22281
R:Almscough, R.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19541
A:Accession: T22281
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-868 <MIL>
A:Cross-references: EMBL:Z01540; PIDN:CA804405.1; GSPDB:GN00023; CESP:F46B3.5
A:Experimental source: clone F46B3
A:Genetics:
A:Gene: CESP:F46B3.5
A:Map position: 5
A:Introns: 45/3; 200/1; 235/3; 436/3; 561/3; 581/3; 696/3; 742/1; 826/3; 846/3

```

Query Match	10.0%	Score 209	DB 2	Length 866
Best Local Similarity	30.4%	Pred. No. 7	5e-09	
Matches	63	Conservative	37	Mismatches 83; Indels 24; Gaps 8
QY	197	GGCGPQNAATVRLMSGCRKRLRELRHSDWLVADASGRVVP--TPVLFLFDRIQIRASTV	254	
Db	662	GGCFSSDILVTTPSG-KRKNDELIDVGVDVLLAN--RVATHTFLPVLIMTHRSEKKEEL	717	

```

0Y 255 AVETWMPRKLLTTPMHLVFEARBPAPA-----PGDFAPYFARRLKAGSYLAPGDAL 308
Db 718 TITTE-RGSTQLTPPLHMYFRKNESESEFKLLIPENHEALALAYLIGCVIITENTKF
0Y 309 RPARVARARBEAGVAFPLAHGTLTYNDVLASCYAVLE-----SHONAHRAFLPLR- 362
Db 777 ROEKINQTTGRKLTGKITISPLKKNRRIIVDMALASCYSEVQANVIQTTTFW---VFNRLRQ 833
0Y 362 -LHHLGALLPGGAVOPTGMHMYSRLL 387
Db 834 KYLNLFGLIHNNHEIPLPTGTVVYELL 860

```

```

RESULT 12
T24045
hypothetical protein R08B4.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T24045
R:White, S.
submitted to the EMBL Data Library, November 1995
A:Reference number: Z19834
A:Accession: T24045
A>Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-1226 <WTL>
A:Cross-references: EMBL,Z68008; PIDN:CAA92000.1; GSPDB:GN00028; CESP:R08B4.1
A:Experimental source: clone R08B4
C:Genetics:
A:Gene: CESP:R08B4.1
A:Map position: 10
A:Intons: 53/2; 113/1; 152/3; 204/3; 272/1; 354/1; 389/3; 714/3; 839/3; 877/1; 960/3

```

[illegible]

```

RESULT 13
T33252
hypothetical protein K02E2.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence
C:Accession: T33252
R:Lloyd, C.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19716
A:Accession: T33252
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1021 <NT>
A:Cross-references: EMBL:Z81560; P1DN:CA04547.1;
A:Experimental source: clone K02E2
A:Gene: CESP:K02E2.2
A:Map position: 5

```

A:introns: 41/1; 79/3; 131/3; 197/1; 262/1; 297/3; 566/3; 610/3; 735/3; 775/3; 834/3; 88

Query Match

9.3%; Score 193.5; DB 2; Length 1021;

Best Local Similarity 27.1%; Pred. No. 1.6e-07;

Matches 59; Conservative 41; Mismatches 83; Indels 35; Gaps 8;

QY 199 CFPGNATVRLMSGERKGLRELHGRDVLADASGRVPPVLLFLDRDQRRASFVAVET 258

DB 802 CFSRDVWVTPSG-KKRMDEIEIGDVLTDLTALF-SAITLMHREPEVQEFLEIKT 859

QY 259 EMPPRKLLTPMHLVFAAR-----GPAP-----ACDPAPYFARLRAG 297

DB 860 D-NGKTLQLAGHFIATPCRLPSKNSSLNSTPERYHLLDTLPDSETKLASQKIG 918

QY 298 DSVLAPGDALPARYARAREAVGFAPLTAGHTLVNDVLAACYAVLESHQ----- 352

DB 919 ECLLHNGQFNMOKIDTSKTVSTIGISPLTENGRIIVNDVLAACYSEVQNVLOTTF 978

QY 352 WAHRAFAPIR--LHAALGALLPGGAVOPTGMHYSRL 387

DB 979 W--AFDRRLNLIQYFGDLYDEIPLPTGSLYKEVL 1013

RESULT 14

123754
hypothetical protein T05C12.10 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999

C:Accession: T23754; T24513

R:Thomas, K.
submitted to the EMBL Data Library, June 1995

A:Reference number: 219793

A:Accession: T23754

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1207 <M12>

A:Cross-references: EMBL:Z49968; PIDN:CAA90265.1; GSPDB:GN00020; CESP:T05C12.10

A:Experimental source: clone M110

R:Burton, J.
submitted to the EMBL Data Library, October 1995

A:Reference number: 219901

A:Accession: T24513

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1207 <M12>

A:Cross-references: EMBL:Z66500; PIDN:CAA91313.1; GSPDB:GN00020; CESP:T05C12.10

A:Experimental source: clone T05C12

C:Genetics:

A:Gene: CESP:T05C12.10

A:Map position: 2

A:introns: 31/3; 87/2; 141/3; 180/2; 203/3; 267/1; 776/2; 794/2; 834/2; 1086/3; 1143/1;

Query Match

8.9%; Score 186.5; DB 2; Length 1207;

Best Local Similarity 31.9%; Pred. No. 7e-07;

Matches 60; Conservative 36; Mismatches 71; Indels 21; Gaps 8;

QY 185 SYKANSLAVRAGG---CFPGNATVRLMSGERKGLRELHGRDVLADASGRVPPV 239

DB 942 AVLATTPGAGAGGGRSNCFSADSLVTTVTOQR-MDELQIGDYLVPPSSGNVLKYEK 1000

QY 240 LLEFLDRDQRRASFVAVETEMPRLKLLTPMHLVFAARGP-----PAPGDFP---PV 289

DB 1001 EMFYHREPRETRNFVLYLK-SGRKLSLGRHLLPYAECSEQVEQYTMNPDGIDVAMRESK 1059

QY 290 FARRLRAGDSVLA--PGGDALPARYARAREAVGFAPLTAGHTLVNDVLAACYAVL 347

DB 1060 YAEKARKGCEVLSIDSEGVAD-EIVRVGMTNVGIYSPMTVEGSLIVDGLVSSCFSHL 1118

QY 348 ESHQWHR 355

DB 1119 ESHS-AHK 1125

RESULT 15

127665
hypothetical protein ZK1037.10 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T27665

R:Basham, V.
submitted to the EMBL Data Library, October 1996

A:Reference number: 220401

A:Accession: T27665

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-481 <M12>

A:Cross-references: EMBL:Z81142; PIDN:CA803509.1; GSPDB:GN00023; CESP:ZK1037.10

A:Experimental source: clone ZK1037

C:Genetics:

A:Gene: CESP:ZK1037.10

A:Map position: 5

A:introns: 37/3; 75/3; 141/3; 173/2; 253/1; 298/3; 310/1; 371/2; 424/3

Query Match

8.7%; Score 182; DB 2; Length 481;

Best Local Similarity 29.6%; Pred. No. 5.2e-07;

Matches 47; Conservative 31; Mismatches 73; Indels 8; Gaps 4;

QY 199 CFPGNATVRLMSGERKGLRELHGRDVLADASGR-VVPTPVLLFLDRDQRRASFVAVET 257

DB 314 CFPNDVAVNTEKAVKRMDELEIGDWEALDENGEDITFLPVKTYLHRDPQEAFFLEFS 373

QY 258 TEMPRLKLLTPMHLVFAA--RGAPAPGDFAPVFAARLRAGDSVLA PGDAL--RPA 311

DB 374 LD-NGEFTLREKHLYVTECRQNSSELKISWESISAGKVNAAGOCFYLAQSEALTKRYLV 432

QY 312 RVARAREAVGFAPLTAGHTLVNDVLAACYAVLESH 350

DB 433 ELDIKRVKKTGIYAPMTSGHLLVNNKIHTSCHSEVDHH 471

Search completed: June 5, 2000, 08:17:55
Job time: 1235 sec

THIS PAGE BLANK (USPTO)

1945

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2000, 08:19:54 ; Search time 14.43 Seconds
(without alignments)
835.771 Million cell updates/sec

Title: US-08-900-220-17

Perfect score: 2088
Sequence: 1 MALLTNLPICCLALLALPA.....PTGMWYSLLYRLAEELLG 396

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 83857 seqs, 30454973 residues

Total number of hits satisfying chosen parameters: 83857

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : SWISSPROT_38:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	4688	100.0	396 1 DHH_HUMAN	043323 homo sapien
2	2008	96.2	396 1 DHH_MOUSE	061488 mus musculu
3	1273.5	61.0	396 1 DHH_XENLA	091610 xenopus lae
4	1255.5	60.1	398 1 DHH2_XENLA	091611 xenopus lae
5	1214	58.1	437 1 SHH_MOUSE	062226 mus musculu
6	1203	57.6	437 1 SHH_RAT	063673 rattus norv
7	1198	57.4	425 1 SHH_CHICK	091035 gallus gall
8	1177	56.4	462 1 SHH_HUMAN	013465 homo sapien
9	1161.5	55.6	411 1 SHH_HUMAN	014623 homo sapien
10	1160	55.6	408 1 SHH_CHICK	098938 gallus gall
11	1154	55.3	411 1 SHH_MOUSE	097812 mus musculu
12	1154	55.3	416 1 SHH_MOUSE	090419 brachydanio
13	1131.5	54.2	432 1 SHH_BRARE	090385 cynops pyr
14	1120	53.6	418 1 SHH_BRARE	092008 brachydanio
15	1092.5	52.3	444 1 SHH_XENLA	092008 xenopus lae
16	1080.5	51.7	409 1 SHH_XENLA	091612 xenopus lae
17	1022	48.9	412 1 SHH_BRARE	098862 brachydanio
18	943	45.2	471 1 SHH_DROME	002936 drosophila
19	883.5	42.3	481 1 SHH_DROME	056574 drosophila
20	484.5	23.2	121 1 SHH_CARAU	P7691 carassius a
21	484.5	23.2	121 1 SHH_PUNTE	P79850 punctius tet
22	482.5	23.1	121 1 SHH_DANAE	013235 danio aff.
23	482.5	23.1	121 1 SHH_DANAE	013234 danio aequi
24	482.5	23.1	121 1 SHH_DANAT	013238 danio aff.
25	482.5	23.1	121 1 SHH_DANAT	013245 danio frank
26	482.5	23.1	121 1 SHH_DANKE	P79709 danio kerri
27	482.5	23.1	121 1 SHH_DANPU	P79717 danio pulch
28	482.5	23.1	121 1 SHH_DEVDE	013241 devario dev
29	482.5	23.1	121 1 SHH_DEVMA	013247 devario mal
30	482.5	23.1	121 1 SHH_DEVPA	013247 devario pat
31	482.5	23.1	121 1 SHH_TENAL	P79915 tenichthys
32	481.5	23.0	121 1 SHH_RASSEL	P79858 rasbora ele
33	480.5	23.0	121 1 SHH_RASPA	P79869 rasbora pav
34	478.5	22.9	121 1 SHH_RASHE	P79864 rasbora het

35	475.5	22.8	121 1 SHH_AMBCH	P79662 amblypharyn
36	475.5	22.8	121 1 SHH_PUNCO	P79838 punctius con
37	291	13.9	58 1 DHH_BRARE	P79729 brachydanio
38	290	13.9	58 1 DHH_DANKE	P79712 danio kerri
39	290	13.9	58 1 SHH_CARAU	P79693 carassius a
40	290	13.9	58 1 SHH_DANAT	013240 danio aff.
41	290	13.9	58 1 SHH_DANKE	P79711 danio kerri
42	290	13.9	58 1 SHH_DANPU	P79719 danio pulch
43	290	13.9	58 1 SHH_DEVDE	013243 devario dev
44	290	13.9	58 1 SHH_PUNTE	P79852 punctius tet
45	290	13.9	58 1 SHH_RASSEL	P79860 rasbora ele

ALIGNMENTS

RESULT 1

ID	DHH_HUMAN	STANDARD	PRT	396 AA.
AC	043323; Q15794;			
DT	15-JUL-1999 (Rel. 38, Created)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	DESERT HEDGEHOG PROTEIN PRECURSOR (DHH) (HMG-3).			
GN	DHH.			
OS	Homo sapiens (Human).			
CC	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;			
CC	Eutheria; Primates; Catarrhini; Homiidae; Homo.			
CC	[1]			
RP	SEQUENCE FROM N.A.			
RA	Tate G., Endo Y., Mitsuura T.;			
RL	Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.			
RL	[2]			
RP	SEQUENCE OF 85-178 FROM N.A.			
RC	TISSUE-KIDNEY;			
RA	Drummond I.A.;			
RT	"Human desert hedgehog.";			
RL	Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.			
CC	- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT. MAY FUNCTION AS A SPERMATOCTYTE SURVIVAL FACTOR IN THE TESTES. ESSENTIAL FOR TESTES DEVELOPMENT.			
CC	- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).			
CC	- PFM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).			
CC	- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.			
CC	- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AB010994; BAA24866.1;			
DR	EMBL; AB010581; BAA24866.1; JOINED.			
DR	EMBL; AB010993; BAA24866.1; JOINED.			
DR	EMBL; U59748; AAB03398.1;			
DR	PFAM; PF01079; Hhnt; 1.			
DR	PFAM; PF01085; Hh-signal; 1.			
DR	PRINTS; PR00632; SONICHOG.			
KW	Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;			

```

KW  signal.
FT  SIGNAL. 1 22
FT  CHAIN 23 396
FT  CHAIN 23 198
FT  CHAIN 199 396
FT  CHAIN 199 396
FT  SITE 198 244
FT  SITE 244 244
FT  SITE 244 244
FT  SITE 268 268
FT  ACT_SITE 271 271
FT  BINDING 198 198
FT  CONFLICT 177 177
SQ  SEQUENCE 396 AA: 43577 MW: FCEAFB21972C3AD5 CRC64:

Query Match 100.0% Score 2088: DB 1: Length 396:
Best Local Similarity 100.0% Pred. No. 1.9e-169;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALTLNLPCLALALPAOSCGPGRGRRRARKQVPLLYKQVPGVPEPTLGAS 60
    |||
DB 1 MALTLNLPCLALALPAOSCGPGRGRRRARKQVPLLYKQVPGVPEPTLGAS 60
    |||

QY 61 GPAEGRVARGSERFDLVNPNYNDIIFKDENSGADRLMTERCKERYNALAIYVNMMPG 120
    |||
DB 61 GPAEGRVARGSERFDLVNPNYNDIIFKDENSGADRLMTERCKERYNALAIYVNMMPG 120
    |||

QY 121 VRLRVTEGMEDEGHHADSLHYEGRALDITTSRDRNRYGLLARLAVEAGFDWVYYESRN 180
    |||
DB 121 VRLRVTEGMEDEGHHADSLHYEGRALDITTSRDRNRYGLLARLAVEAGFDWVYYESRN 180
    |||

QY 181 HHVSVKADNSLAVRAGGCPGNATVRLMSGERRGLRELHSGDVLAAASGRVVTPLY 240
    |||
DB 181 HHVSVKADNSLAVRAGGCPGNATVRLMSGERRGLRELHSGDVLAAASGRVVTPLY 240
    |||

QY 181 HHVSVKADNSLAVRAGGCPGNATVRLMSGERRGLRELHSGDVLAAASGRVVTPLY 240
    |||
DB 181 HHVSVKADNSLAVRAGGCPGNATVRLMSGERRGLRELHSGDVLAAASGRVVTPLY 240
    |||

QY 241 LFLDRDLQRRASVAVETEMPRKLLTPMHLFAARGPARADPAFARLRAGDSY 300
    |||
DB 241 LFLDRDLQRRASVAVETEMPRKLLTPMHLFAARGPARADPAFARLRAGDSY 300
    |||

QY 301 LAGGDLARVARARARERAVFPALTHGTLVNDVLAASYAVLESQVHRAFAPL 360
    |||
DB 301 LAGGDLARVARARARERAVFPALTHGTLVNDVLAASYAVLESQVHRAFAPL 360
    |||

QY 361 RLHLALGALLPGGAVOPTGMHWSRLLYRLAEELG 396
    |||
DB 361 RLHLALGALLPGGAVOPTGMHWSRLLYRLAEELG 396
    |||

RESULT 2
DHH_MOUSE
ID DHH_MOUSE STANDARD: PRT: 396 AA.
AC 061488:
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE DESERT HEDGEHOG PROTEIN PRECURSOR (DHH) (HMG-3).
GN DHH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN 11)
RP SEQUENCE FROM N.A.
RC STRAIN-129/SV;
RX MEDLINE: 94094334.
RA Echelard Y., Epstein D.J., St Jacques B., Shen L., Mohler J.,
RA McMahon J.A., McMahon A.P.;
RT "Sonic hedgehog, a member of a family of putative signaling
RT molecules, is implicated in the regulation of CNS polarity.";
RL Cell 75:1417-1430(1993).
RP 12)
RP SEQUENCE OF 120-168 FROM N.A.

```

```

RX MEDLINE: 95236997.
RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,
RA Zhao R., Seldin M.F., Fallon J.F., Beachy P.A.;
RT "Products, genetic linkage and limb patterning activity of a murine
RT hedgehog gene.";
RL Development 120:339-353(1994).
CC -!- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
CC PATTERNING EVENTS DURING DEVELOPMENT. MAY FUNCTION AS A
CC SPERMATOCTE SURVIVAL FACTOR IN THE TESTES. ESSENTIAL FOR TESTES
CC DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
CC CELL SURFACE. WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
CC CELL SURFACE, IT IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ADULT TESTES. NOT EXPRESSED IN
CC LIMB BUDS.
CC -!- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X76292; CAA53924.1; -.
DR MGD: MGI:94891; DHH.
DR PFAM: PF01079; Hntc; 1.
DR PFAM: PF01085; HH_signal; 1.
DR PRINTS: PR00632; SONICHHOG.
KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
KW signal.
FT SIGNAL. 1 22
FT CHAIN 23 396
FT CHAIN 23 198
FT CHAIN 199 396
FT CHAIN 199 396
FT SITE 244 244
FT SITE 244 244
FT SITE 268 268
FT ACT_SITE 271 271
FT BINDING 198 198
FT CONFLICT 177 177
SQ SEQUENCE 396 AA: 43542 MW: AFEB051BE950FD8 CRC64:

Query Match 96.2% Score 2088: DB 1: Length 396:
Best Local Similarity 96.5% Pred. No. 1.1e-162;
Matches 382; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 MALTLNLPCLALALPAOSCGPGRGRRRARKQVPLLYKQVPGVPEPTLGAS 60
    |||
DB 1 MALTLNLPCLALALPAOSCGPGRGRRRARKQVPLLYKQVPGVPEPTLGAS 60
    |||

QY 61 GPAEGRVARGSERFDLVNPNYNDIIFKDENSGADRLMTERCKERYNALAIYVNMMPG 120
    |||
DB 61 GPAEGRVARGSERFDLVNPNYNDIIFKDENSGADRLMTERCKERYNALAIYVNMMPG 120
    |||

QY 121 VRLRVTEGMEDEGHHADSLHYEGRALDITTSRDRNRYGLLARLAVEAGFDWVYYESRN 180
    |||
DB 121 VRLRVTEGMEDEGHHADSLHYEGRALDITTSRDRNRYGLLARLAVEAGFDWVYYESRN 180
    |||

```

QY 181 HHVSVKADNSLAVRAGGCPGNATVRLMSGKRLRELRHGDVLAADAGRVVPPVL 240
 DB 181 HHVSVKADNSLAVRAGGCPGNATVRLMSGKRLRELRHGDVLAADAGRVVPPVL 240
 QY 241 LFLDRDQRRASFVAETEMPPKLLTPMHLVFAARGPAPAGDFAPVFAARRLRAGDSV 300
 DB 241 LFLDRDQRRASFVAETEMPPKLLTPMHLVFAARGPAPAGDFAPVFAARRLRAGDSV 300
 QY 301 LAPGDAALRPARVARVAREAVGAPLTAHGTLLVNDVLASCAYAVLESHQMAHRAFAPL 360
 DB 301 LAPGDAALRPARVARVAREAVGAPLTAHGTLLVNDVLASCAYAVLESHQMAHRAFAPL 360
 QY 361 RLHLALGALLPGAGVOTGMWYSRLYLRLAEELG 396
 DB 361 RLHLALGALLPGAGVOTGMWYSRLYLRLAEELG 396
 RESULT 3
 DH1_XENLA STANDARD; PRT; 396 AA.
 AC Q91610;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE DESERT HEDGEHOG PROTEIN PRECURSOR 1 (DHH-1) (CEPHALIC HEDGEHOG PROTEIN) (X-CHH).
 GN CHH.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Amphibia;
 OC Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae; Xenopodinae;
 OC Xenopus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-EMBRYO;
 RX MEDLINE: 95401852.
 RA Ecker S.C., McGrew L.L., Lai C.-J., Lee J.J., von Kessler D.P., Moon R.T., Beachy P.A.;
 RT "Distinct expression and shared activities of members of the hedgehog gene family of Xenopus laevis.";
 RL Development 121:2337-2347(1995).
 CC -1- FUNCTION: SIGNAL INVOLVED IN THE EARLY INDUCTION AND PATTERNING OF ANTERODORSAL ECTODERM, NERVOUS SYSTEM AND SOMITES. INDUCES ECTOPIC CEMENT GLAND FORMATION IN EMBRYOS.
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE CELL, WHILE THE C-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE CELL SURFACE. HEDGEHOG PROTEIN IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- DEVELOPMENTAL STAGE: DETECTABLE WITHIN THE EARLY GASTRULA. IN THE NEURULA EXPRESSION BECOMES RESTRICTED TO ANTERIOR STRUCTURES, ENCOMPASSING BOTH NEURAL PLATE AND ENDODERMAL CELLS.
 CC -1- PFM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: U26349; AAA85163.1;
 DR PFAM: PF01079; Hntc; 1.
 DR PFAM: PF01085; HH_signal; 1.

DR PRINTS: PRO0632; SONTCHOG.
 KW Developmental Protein; Autocatalytic cleavage; Hydrolase; Protease;
 KM Signal.
 FT SIGNAL 1 22
 FT CHAIN 23 396
 FT CHAIN 23 396
 FT CHAIN 198 396
 FT SITE 197 198
 FT SITE 267 267
 FT ACT_SITE 270 270
 FT BINDING 197 197
 FT DOMAIN 276 279
 SQ SEQUENCE 396 AA; 44087 MW; 774A3EC2268A5EE9 CRC64;
 Query Match 61.0%; Score 1273.5; DB 1; Length 396;
 Best Local Similarity 63.6%; Pred. No. 1.7e-100;
 Matches 248; Conservative 52; Mismatches 83; Indels 7; Gaps 5;
 QY 7 LPLPCTALLALPAOSGCPGRRYARKOLVPLLYKOFVGPVETIGASGPAEGR 66
 DB 8 ILAICG-GLLVPRCCGPGRRYARKOLVPLLYKOFVGPVETIGASGPAEGR 65
 QY 67 VARGSEFRDLPVNPDIIFKDENSGADRLMTERCKERYNALAIAMNMPGVRLVLT 126
 DB 66 IHRSERIEIPLVNPDIIFKDEKQADRLMTERCKERYNALAIAMNMPGVRLVLT 125
 QY 127 EGMWEDHHAODSLHYGRALDITTSRDRNKYGLRLAVLAEAGFDWYIESRHHVSV 186
 DB 126 EGMWEDHHAODSLHYGRALDITTSRDRNKYGLRLAVLAEAGFDWYIESRHHVSV 185
 QY 187 KADNSLAVRAGGCPGNATVRLMSGKRLRELRHGDVLAADAGRVVPPVLFLDRD 246
 DB 186 KADNSLAVRAGGCPGNATVRLMSGKRLRELRHGDVLAADAGRVVPPVLFLDRD 245
 QY 247 LQRRASEFAVATEMPKLLTPMHLVFAARGPAPAGDFAPVFAARRLRAGDSV-LAPGG 305
 DB 246 PYKATFVLEAEHGPRLVTPHLLFI--QSSSAGFLPTFAVYQIGDVIQVYNG 302
 QY 306 DALPAPARVARVAREAVGAPLTAHGTLLVNDVLASCAYAVLESHQMAHRAFAPLRLHA 365
 DB 303 TQVSSKVVYRSLSEOTGVAPMTEHGTLLVNDVLASCAYAVLESHQMAHRAFAPLRLFOG 362
 QY 366 LGALLPGAGVOTGMWYSRLYLRLAEEL 395
 DB 363 TASMIP-DLDMSDGVHWCILYLAARYL 391
 RESULT 4
 DH2_XENLA STANDARD; PRT; 398 AA.
 AC Q91611;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE DESERT HEDGEHOG PROTEIN PRECURSOR 2 (DHH-2) (HEDGEHOG PROTEIN 4) (X-HH).
 GN HH4.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Amphibia;
 OC Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae; Xenopodinae;
 OC Xenopus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-EMBRYO;
 RX MEDLINE: 95401852.
 RA Ecker S.C., McGrew L.L., Lai C.-J., Lee J.J., von Kessler D.P., Moon R.T., Beachy P.A.;
 RT "Distinct expression and shared activities of members of the hedgehog gene family of Xenopus laevis.";
 RL Development 121:2337-2347(1995).

[illegible]

```

DB 304 GTVOSSKVVRSVDEQGTGYVAPMTPEHGTLVDGVLTSCYATVESHTLAA5LAPLRIFQ 363
OY 365 ALGALLPGAVOPTGMHWYSRLLYRAEELL 395
Db 364 GIASMLP-DLHTSDGVHWYCHILLYLAKYVL 393

RESULT 5
SHH_MOUSE
ID SHH_MOUSE STANDARD: PRT: 437 AA.
AC 062226;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE SHH OR HEDGEHOG PROTEIN PRECURSOR (SHH) (HMG-1).
CN SHH OR HHGL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RA Echelard Y., Epstein D.J., St Jacques B., Shen L., Mohler J.,
RA McMahon J.A., McMahon A.P.;
RT "Sonic hedgehog, a member of a family of putative signaling
RT molecules, is implicated in the regulation of CNS polarity.";
RL Cell 75:1417-1430(1993).
RN [2]
RN REVISION TO 122.
RP STRAIN=C57BL/6J;
RC McMahon A.P.;
RA Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.
RP MEDLINE: 95234697.
RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandi B.K.,
RA Zhao R., Seidlin M.F., Fallon J.F., Beachy P.A.;
RT "Products, genetic linkage and limb patterning activity of a murine
RT hedgehog gene.";
RL Development 120:3339-3353(1994).
RN [4]
RN FUNCTION, AND AUTOPROTEOLYTIC CLEAVAGE.
RP MEDLINE: 95254654.
RA Roelink H., Porter J.A., Chiang C., Tanabe Y., Chang D.T.,
RA Beachy P.A., Jessell T.M.;
RT "Floor plate and motor neuron induction by different concentrations of
RT the amino-terminal cleavage product of sonic hedgehog
RT autoproteolysis.";
RL Cell 81:445-455(1995).
RN [5]
RN X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 34-195.
RX Hall T.M.T., Porter J.A., Beachy P.A., Leahy D.J.;
RT "A potential catalytic site revealed by the 1.7-A crystal structure of
RT the amino-terminal signalling domain of Sonic hedgehog.";
RL Nature 378:212-216(1995).
RN [6]
RN FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN
RN ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION
RN OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE
RN CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER
RN TARGET, THE GGL ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A
RN VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED
RN BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE
RN AND SOMITES. AND THE POLARIZING SIGNAL FOR PATTERNING OF THE
RN ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BOD. DISPLAYS BOTH
RN FLOOR PLATE-AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD
RN CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS
RN 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLANE INDUCTION (BY
RN SIMILARITY)
RN [7]
RN SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
RN CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
RN CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM

```


FT ACT_SITE 273 273 ESSENTIAL FOR AUTO-CLEAVAGE (BY
 FT BINDING 200 200 SIMILARITY).
 FT DOMAIN 390 393 CHOLESTEROL (BY SIMILARITY).
 SO SEQUENCE 425 AA; 46474 MW; DA9627443DA0173 CRC64;

Query Match 57.4%: Score 1198; DB 1; Length 425;
 Best Local Similarity 59.4%: Pred. No. 4.4e-94;
 Matches 244; Conservative 56; Mismatches 91; Indels 20; Gaps 10;

QY 1 MALLTNLP---LCCLALLAPASCGGRRPVRARRARQVPLLYKQVPPVPEPTL 57
 D 4 MLLTRILLVGFIC--ALVSSGLTCGGNG-ICKRRPK-LTPRLAKQEPFNPVAKTLL 59
 QY 58 GASPAAGRVARGSEFPDLPVNPPIIFDEENSGADRMTRECKRVNLAIVANN 117
 D 60 GASGRYRGKTRNSERKELTPNNPDIIFDEENSGADRMTRECKRVNLAIVANN 119
 QY 118 WPGVRLNTEGMDGHHADSLHSEGRALDITTSRDNRNKGGLARLAVEGFDWYEE 177
 D 120 WPGVRLNTEGMDGHHSESLHYEGRAVDITTSRDNRNKGGLARLAVEGFDWYEE 179
 QY 178 SRNVHVSVDNLSNLRVAGGCFPGNATVRLMSGERKGLRLHGDVYLADASGRVPT 237
 D 180 SKAHICSVKAEVNSVAAKSGGCFPGSATVHLHGGTKLVKDLSPDRYLADADGRLLYS 239
 QY 238 PVLFLDRLDRASFVAVERTEPPKRLLETPMHLVFA--RGPAAPAGDFA--PVEARR 293
 D 240 DFLFLDRMSSRLFLYIETROPARRALLTAHLFLFAAPHNOSSEATGSGALFPMN 299
 QY 294 LRAGDS--VIAFGDALPAPAVANA--REAVGVAPFLTAGTLVNDVLASCYAVLESH 350
 D 300 VKPQRYVYLGEGGQQLLPASVHSVSLKEASGAVAPFLTAGTLVNDVLASCYAVIEH 359
 QY 351 QMARAPAPLRLAHLA-LGALLPGAV-----QPTGMHYSLLYRLABELL 395
 D 360 SWAHMAFPRLAGLLAALCPDAIPTAATTGTHIHSYRLYIGSWL 410

RESULT 8
 SHH_HUMAN STANDARD; PRT; 462 AA.
 AC Q15465;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE SONIC HEDGEHOG PROTEIN PRECURSOR (SHH) (HNG-1).
 GN SHH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE: FETAL LUNG;
 RX MEDLINE; 96070431.
 RA Marigo V., Roberts D., Lee S.M.K., Tsukurov O., Levi T.,
 RA Gaslier J.M., Epstein D.J., Gilbert D.J., Copeland N.G., Seidman C.E.,
 RA Jenkins N.A., Seidman J.G., McMahon A.P., Tabin C.;
 RT "Cloning, expression, and chromosomal location of SHH and IHH: two
 RT human homologues of the Drosophila segment polarity gene hedgehog.";
 RL Genomics 28:44-51(1995).
 RN [2]
 RP SEQUENCE OF 1-187 FROM N.A.
 RA Strong C., Graves T., Suterer C., Ozersky P.;
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 119-167 FROM N.A.
 RX MEDLINE; 95236997.
 RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Samandl B.K.,
 RA Zhao R., Seidlin M.F., Fallon J.F., Beachy P.A.;
 RT "Products, genetic linkage and limb patterning activity of a murine
 hedgehog gene.";

RL Development 120:3339-3353(1994).
 RN [4]
 RP VARIANTS HPE3 ARG-31, GLY-117 AND ARG-117.
 RX MEDLINE; 97051937.
 RA Roessler E., Belloni E., Gaudenz K., Jay P., Berta P., Scherer S.W.,
 RA Tsui L.-C., Muenke M.;
 RT "Mutations in the human Sonic Hedgehog gene cause holoprosencephaly.";
 RL Nat. Genet. 14:357-360(1996).
 CC -1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN
 CC ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION
 CC OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE
 CC CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER
 CC TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A
 CC VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED
 CC BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE
 CC AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE
 CC ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH
 CC FLOOR PLATE AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD
 CC CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS
 CC 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN FETAL INTESTINE, LIVER, LUNG, AND
 CC KIDNEY. NOT EXPRESSED IN ADULT TISSUES.
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEMLY*GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- DISEASE: DEFECTS IN SHH ARE THE CAUSE OF THE AUTOSOMAL DOMINANT
 CC DISORDER HOLOPROSENCEPHALY TYPE 3 (HPE3). HPE3 IS A GENETICALLY
 CC HETEROGENEOUS DISEASE THAT AFFECTS THE MIDLINE DEVELOPMENT OF THE
 CC FOREBRAIN AND MIDFACE. HPE IS ASSOCIATED WITH SEVERAL DISTINCT
 CC FACIES AND PHENOTYPIC VARIABILITY. IN THE MOST EXTREME CASES,
 CC ANOPHTHALMIA OR CYCLOPIA IS EVIDENT ALONG WITH A CONGENITAL
 CC ABSENCE OF THE MATURE NOSE. THE LESS SEVERE FORM FEATURES FACIAL
 CC DYSMORPHIA CHARACTERIZED BY OCULAR HYPERTELEORISM, DEFECTS OF THE
 CC UPPER LIP AND/OR NOSE, AND ABSENCE OF THE OLFACTORY NERVES OR
 CC CORPUS CALLOSUM.
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L38518; AAA62179.1; -
 CC EMBL; AC002484; AAB67604.1; -
 CC MIM; 600725; -
 CC MIM; 142945; -
 CC DR PFMW; PFO1079; Hnt; 1.
 CC DR PFMW; PFO1085; HH_signal; 1.
 CC DR PRINTS; PRO0632; SONICHOG.
 CC Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 CC Signal; Disease mutation.
 CC FT SIGNAL 1 23
 CC FT CHAIN 24 462
 CC FT CHAIN 24 197
 CC FT CHAIN 198 462
 CC FT SITE 197 198
 CC FT SITE 243 243
 CC FT SITE 267 267
 CC INVOLVED IN AUTO-CLEAVAGE (BY
 CC SIMILARITY).
 CC INVOLVED IN CHOLESTEROL TRANSFER (BY
 CC SIMILARITY).
 CC INVOLVED IN AUTO-CLEAVAGE (BY

```

FT ACT_SITE 270 270 SIMILARITY)
FT BINDING 197 197 ESSENTIAL FOR AUTO-CLEAVAGE (BY
FT DOMAIN 407 411 SIMILARITY).
FT VARIANT 31 31 CHOLESTEROL (BY SIMILARITY).
FT VARIANT 117 117 G -> R (IN HPE3).
FT VARIANT 117 117 W -> G (IN HPE3).
FT VARIANT 117 117 /FTID-VAR_003620.
FT VARIANT 117 117 W -> R (IN HPE3).
FT VARIANT 117 117 /FTID-VAR_003621.
FT VARIANT 117 117 /FTID-VAR_003621.
SO SEQUENCE 462 AA; 49607 MW; D687AFAS82A4749 CRC64;

Query Match 56.4%; Score 1177; DB 1; Length 462;
Best Local Similarity 54.4%; Pred. No. 3e-92;
Matches 245; Conservative 54; Mismatches 93; Indels 58; Gaps 10;

QY 1 MALLTNLPDLCL-ALLAPASCGRGPRVARRARQVPLLYKQVPGVPETLGA 59
DB 1 MLLARCLLVLSLVSCGLACGGRG-FGKRHPK-KLPPVAKOPIPVNAERTLGA 58
QY 60 SGPASGRVARGSERFDVLPNTNPDIIFDENSGADRLMTERCKERVVALAVANNMP 119
DB 59 SGREYEGKIRNSERREKELTPNPNPDIIFDENTGADRLMTORCKDKLVALISVANNMP 118
QY 120 GVRNLTVEGMDGHHADSLHYEGRALDITTSDRNRKYLGLARLAVAGEFWYTESR 179
DB 119 GVKRLTEGMDGHHSESLHYEGRAVDITTSDRNRKYLGLARLAVAGEFWYTESR 178
QY 180 NHVHVSVDNLSAVRAGCGFPGNATVRLMSEKGLRLHGDWVLADAGRVVPY 239
DB 179 AIHOSVAVNSVAAKSGCGFPGSATVHEOGKTLVLDLSPGDVLAADDGRLLSDF 238
QY 240 LFLFDRLORASFVAVEIEMPPKLLTPMHLVFAA-----RGAPAPAG 284
DB 239 LFLFDRLORASFVAVEIEMPPKLLTPMHLVFAA-----RGAPAPAG 284
QY 285 DEAP--VFARRLRADDSVLA---PGDALRPARAVARA-REAVGVPAFLTHAGTLVN 337
DB 298 ALGPRLAFASRVPCQRYVVAERDGRLLPRAVHVSYLSEAGAVAPLTAQSTILIN 357
QY 338 DVLASCVAVLESHOMAHAFARLLHNLGALL----- 371
DB 358 RVLASCVAVLESHOMAHAFARLLHNLGALL----- 371
QY 371 PGGAIVP-----TGMHWYSRLLYRLAEELL 395
DB 418 PGADAPAGATAGIHWISQLLYOIGTWLL 447

RESULT 9
ID IHH_HUMAN STANDARD; PRT; 411 AA.
AC Q14623; O43322;
DT 15-JUL-1989 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE INDIAN HEDGEHOG PROTEIN PRECURSOR (IHH) (HMG-2).
GN IHH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominiidae; Homo.
RN RP SEQUENCE FROM N.A.
RA Tate G., Endo Y., Mitsuya T.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN RP SEQUENCE OF 100-411 FROM N.A.
RC TISSUE=FERAL LUNG;
RX MEDLINE; 96070431.
RA Marigo V., Roberts D.J., Lee S.M.K., Tsukurov O., Levi T.,

```

```

RA Gastler J.M., Epstein D.J., Gilbert D.J., Copeland N.G., Seldman C.E.,
RA Jenkins N.A., Seldman J.G., McMahon A.P., Tabin C.;
RT "Cloning, expression, and chromosomal location of SHH and IHH: two
RT human homologs of the Drosophila segment polarity gene hedgehog.";
RL Genomics 28:44-51(1995).
RN [3]
RP SEQUENCE OF 124-172 FROM N.A.
RX MEDLINE; 95236997.
RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,
RA Zhao R., Seldin M.F., Fallon J.F., Beachy P.A.;
RT "Products, genetic linkage and limb patterning activity of a murine
RT hedgehog gene".
RL Development 120:3339-3353(1994).
CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
CC PATTERNING EVENTS DURING DEVELOPMENT. BINDS TO THE PATCHED (PTC)
CC RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO
CC ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IMPLICATED IN
CC ENDODERMAL OSSIFICATION. MAY REGULATE THE BALANCE BETWEEN GROWTH
CC AND OSSIFICATION OF THE DEVELOPING BONES. INDUCES THE EXPRESSION
CC OF PARATHYROID HORMONE-RELATED PROTEIN (PTHrP) (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN EMBRYONIC LUNG, AND IN ADULT
CC KIDNEY AND LIVER.
CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB018076; BAA33523.1;
DR EMBL; AB018075; BAA33523.1; JOINED.
DR EMBL; AB010092; BAA33523.1; JOINED.
DR EMBL; L38517; AAA62178.1;
DR MIM; 600726;
DR PFM; PF01079; Hnt; 1.
DR PFM; PF01085; HH.signal; 1.
KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
KW signal.
FT SIGNAL 1 27
FT CHAIN 28 411
FT CHAIN 28 202
FT CHAIN 203 411
FT SITE 202 203
FT SITE 248 248
FT SITE 272 272
FT ACT_SITE 275 275
FT BINDING 202 202
FT CONFLICT 100 100
FT CONFLICT 244 246
FT CONFLICT 259 259
SO SEQUENCE 411 AA; 45264 MW; 788BAC6B8C0B5647 CRC64;

```

Query Match 55.6%; Score 1161.5; DB 1; Length 411;

[illegible]

CC	C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
CC	-I- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; U58511; AAC60010.1; -
DR	PFAM; PF01079; Hint; 1.
DR	PFAM; PF01085; HH.signal; 1.
DR	PRINTS; PR00532; SONICHOG.
DR	Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
KW	Signal.
FT	1 23
FT	CHAIN 24 408
FT	CHAIN 24 198
FT	CHAIN 199 408
FT	SIZE 198 199
FT	SIZE 244 244
FT	SIZE 268 268
FT	ACT_SITE 271 271
FT	BINDING 198 198
FT	SEQUENCE 408 AA; 44829 MW; BA397AE2A9357A24 CRC64;
SO	
Query Match	55.6%; Score 1160; DB 1; Length 408;
Best Local Similarity	58.7%; Pred. No. 6.9e-91;
Matches 232; Conservative	50; Mismatches 107; Indels 6; Gaps 4;
QY	2 ALTNLPPLCCALALALAAQSGCPGRCVGRRRYARKQLVPLLYKQFYPGVDEFTLGASG 61
DB	4 ALLLLLSGCAI-LLAAPVRCGCGPR-VVGSRRPRLPLIPLAYKQFSPNVEKTLGASG 61
QY	62 PAEGHVAQSGSEFRDLVNVNPDITFKDEBNSGADRLTERCKEKNVNLAIIVMMMPGV 121
DB	62 RREGIARNSERFKELTIRYNNDITFKDEBNSGADRLTERCKEKNVNLAIIVMMMPGV 121
QY	122 RLRTVEGMDDECHNAQDSLHEGRALDITTSDDRNKYGGLARLAVAEQFDDVYYESRNH 181
DB	122 KLRVTEGMDDECHNSESLHTEGRAVDITTSDDRNKYGGLARLAVAEQFDDVYYESKHA 181
QY	182 VAVSVKADNSLAVRAGGCPFGPNATYRLMSGEKGLRELHRGCVYLAAADSGNVPTPYLL 241
DB	182 IHCVSSEHSAAKQGGCPFGPRALTLLEGARTPMALRPPGVRVLAAMPAGAPRTYSDFLA 241
QY	242 FLDRLORASPAVAETEMPRLKLTTPMHLFEAARGPAPAGDEPAFARLRLRAGDSYL 301
DB	242 FLDKEPRALTAHVETROPPRRLLTTPHLLFTVADNASAPAAQRPPTFASVQGHFVL 301
QY	302 -APGGDALRPAFAVARAREAVGAPRLTAHGTLLVNDVLASCVALLSESHQAHRAFAPL 360
DB	302 VAVGGGGLQPAFVGVGRGTIDYGAVAPLTHGTIVLVDDVAVSCFALVRQQLAQAAMFPL 361
QY	361 RLHLHALGALLPGGAVOPTGMHYSKLLRLAEELL 395
DB	362 RLYH--SLTGGPGVQGDGVHWSGLLRYLGRMLL 393
RESULT	11
IHH_MOUSE	
ID	IHH_MOUSE STANDARD; PRT; 411 AA.
AC	P97812; O61724;
DT	15-JUL-1999 (Rel. 38, Created)
DT	15-JUL-1999 (Rel. 38, Last sequence update)
DT	15-JUL-1999 (Rel. 38, Last annotation update)
DE	INDIAN HEDGEHOG PROTEIN PRECURSOR (IHH) (HHG-2).

GN IHH.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.
 RC STRAIN-CD-1; TISSUE-KIDNEY;
 RX MEDLINE: 97236802.
 RA Valentini R.P., Brookhiser W.T., Park J., Yang T., Briggs J.,
 RA Dressler G., Holzman L.B.;
 RT "Post-translational processing and renal expression of mouse Indian
 RT hedgehog.";
 RT J. Biol. Chem. 272:8466-8473(1997).
 RN [2]
 RP SEQUENCE OF 76-411 FROM N.A.
 RC STRAIN-C57BL/6J; TISSUE-EMBRYO;
 RX MEDLINE: 94094334.
 RA Echelard Y., Epstein D.J., St Jacques B., Shen L., Mohler J.,
 RA McMahon A.P., McMahon A.P.;
 RT "Sonic hedgehog, a member of a family of putative signaling
 RT molecules, is implicated in the regulation of CNS polarity.";
 RT Cell 75:1417-1430(1993).
 RN [3]
 RP REVISIONS.
 RC STRAIN-C57BL/6J;
 RA St Jacques B.;
 RT Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 124-172 FROM N.A.
 RX MEDLINE: 95236997.
 RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,
 RA Zhao R., Seldin M.F., Fallon J.F., Beachy P.A.;
 RT "Products, genetic linkage and limb patterning activity of a murine
 RT hedgehog gene.";
 RT Development 120:3339-3353(1994).
 RN [5]
 RP FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 RP PATTERNING EVENTS DURING DEVELOPMENT. BINDS TO THE PATCHED (PTC)
 RP RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHED (SMO), TO
 RP ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IMPLICATED IN
 RP ENDODERMAL OSSIFICATION: MAY REGULATE THE BALANCE BETWEEN GROWTH
 RP AND OSSIFICATION OF THE DEVELOPING BONES. INDUCES THE EXPRESSION
 RP OF PARATHYROID HORMONE-RELATED PROTEIN (PTHrP)
 RP OF PARATHYROID HORMONE-RELATED PROTEIN (PTHrP)
 CC [1] SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC [2] TISSUE SPECIFICITY: IN THE EMBRYO, DETECTED IN THE DEVELOPING GUT,
 CC THE GROWTH ZONE OF CARTILAGE OF DEVELOPING LONG BONES, EPITHELIUM
 CC AND UROGENITAL SINUS. IN THE ADULT KIDNEY, FOUND IN PROXIMAL
 CC CONVOLUTED AND PROXIMAL STRAIGHT TUBULE.
 CC [3] DEVELOPMENTAL STAGE: DETECTED AT 10 DAYS POST COITUM (DPC) IN
 CC IN THE DEVELOPING UROGENITAL SINUS. EXPRESSION INCREASES WITH
 CC GESTATIONAL AGE IN KIDNEY AND DUODENUM, BECOMING MAXIMAL IN
 CC ADULTHOOD.
 CC [4] PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC [5] SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC [6] This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sb.ch).

CC -----
 DR EMBL: U85610; AAB49692.1; ALT_INIT.
 DR EMBL: X76291; CAA53923.1; -
 DR MGI: 96533; IHH.
 DR PFAM: PF01079; Hint: 1.
 DR PFAM: PF01085; HH_Signal: 1.
 DR PRINTS: PR00632; SONCHHOG.
 KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 KW Signal.
 FT SIGNAL 1 27 POTENTIAL.
 FT CHAIN 28 411 INDIAN HEDGEHOG PROTEIN.
 FT CHAIN 28 202 INDIAN HEDGEHOG PROTEIN N-PRODUCT.
 FT CHAIN 203 411 INDIAN HEDGEHOG PROTEIN C-PRODUCT.
 FT SITE 202 203 CLEAVAGE (AUTO-).
 FT SITE 248 248 INVOLVED IN CHOLESTEROL TRANSFER (BY
 FT SITE 248 248 SIMILARITY).
 FT SITE 272 272 INVOLVED IN AUTO-CLEAVAGE (BY
 FT ACT_SITE 275 275 SIMILARITY).
 FT BINDING 202 202 ESSENTIAL FOR AUTO-CLEAVAGE (BY
 FT CONFLICT 383 383 CHOLESTEROL (BY SIMILARITY).
 FT FT W -> S (IN REF. 2).
 SQ SEQUENCE 411 AA: 45485 MM: 08BF7AD8507C0D9B CRC64:
 Query Match 55.3%; Score 1154; DB 1; Length 411;
 Best Local Similarity 59.0%; Pred. No. 2,3e-90;
 Matches 230; Conservative 49; Mismatches 99; Indels 12; Gaps 4;
 QY 12 CLALLAL---PAQSGCPGPGVGRRRYARKQVLPILYKQFVPGVPERLTGASGPAEGRV 67
 DB 13 CFFLLLLVLPARGGCGPGR-VVGSRRPRKVLPLAYKQFPNPKETLGASGREGRTI 71
 QY 68 ARGSEFRDLVNYNDITFDKDENSGADRLMTERTCKEKNALATAYMMMPGVRLRTE 127
 DB 72 ANSEFRKLTPNYNDITFDKDENSGADRLMTERTCKEKNALATAYMMMPGVRLRTE 131
 QY 128 GWDEGHHADSLHVEGRALDITTSDBDRNKYGLRLAVEAGFPMVYVESRHHVSVK 187
 DB 132 GWDEGHHSESLHVEGRALDITTSDBDRNKYGLRLAVEAGFPMVYVESRHHVSVK 191
 QY 188 ADNSLAVRAGCGPFGNATVRLMSGERKGLRELRGDDVLAADASGRVFTPLFLDRDL 247
 DB 192 SEHSAAKTGCGFPAGQVRLNGSERVALSAVPGDRLVAMGEDTPTFSVDLIFLDRDP 251
 QY 248 ORRASVAVETMPRPKLLTPMHLVFAARGAPAPGDPAPFARLRAGDSVLAPGDA 307
 DB 252 NRRAQVETEDPPRRLLTPMHLVFAARGAPAPGDPAPFARLRAGDSVLAPGDA 311
 QY 308 LRPARYARAREAVGFAPLTAHGLTLVNDVLAACYAVLSEHONAHRAFAPLRLHAL- 367
 DB 312 LDPARYAVSTHVALGSYAPLTRHGLVEDVVAACFAVADHLLAQLAFWPLRLFPSLA 371
 QY 367 -GALPAGAVOPTGMHRYRLYLRLAEELL 395
 DB 372 WGSWTPS-----EGVHWYQMLYRLGLRL 396
 RESULT 12
 TWHH_BRARE
 ID TWHH_BRARE STANDARD; PRT; 416 AA.
 AC Q90419;
 DT 15-JUL-1999 (rel. 38, Created)
 DT 15-JUL-1999 (rel. 38, Last sequence update)
 DE 15-FEB-2000 (rel. 39, Last annotation update)
 DE TIGGY-WINKLE HEDGEHOG PROTEIN PRECURSOR (TWHH).
 GN TWHH.
 OS Brachydanio rerio (zebrafish) (zebra danio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
 OC Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Cyprinidae; Rasbora; Danio.
 RN [1]
 RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.

CC TISSUE-EMBRYO:
 CC MEDLINE: 96014264. ●
 RA Ekker S.C., Ungar A.R., Greenstein P., von Kessler D.P.,
 RA Porter J.A., Moon R.T., Beachy P.A.;
 RT "Patterning activities of vertebrate hedgehog proteins in the
 RT developing eye and brain."
 RL Curr. Biol. 5:944-955(1995).
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT. INVOLVED IN DORSO-VENTRAL
 CC PATTERNING OF THE BRAIN AND IN EARLY PATTERNING OF THE DEVELOPING
 CC EYES.
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE VENTRAL MIDLINE OF THE NEURAL
 CC TUBE AND BRAIN. IN THE DEVELOPING BRAIN, EXPRESSION OCCURS IN
 CC DOMAINS THAT INCLUDE A DISCRETE REGION IN THE FLOOR OF THE
 CC Diencephalon. NOT DETECTED IN THE NOTOCHORD OR DEVELOPING FIN BUD.
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to [license@isb-sid.ch](mailto:license@isb-sib.ch)).
 CC
 CC EMBL: U30710; AAC59741.1; -
 CC ZFIN: ZDB-GENE-980526-41; TWHH.
 CC PFAM: PF01079; Hint: 1.
 CC PFAM: PF01085; HH_Signal: 1.
 CC PRINTS: PRO0633; SONICHHOG.
 CC Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 CC Signal.
 CC
 CC FT SIGNAL 1 26
 CC FT CHAIN 27 416
 CC FT CHAIN 27 200
 CC FT CHAIN 201 416
 CC FT SITE 200 201
 CC FT SITE 270 270
 CC FT ACT_SITE 273 273
 CC FT BINDING 200 200
 CC FT SEQUENCE 416 AA; 46576 MW; 61EC2218309CFE59 CRC64;
 CC
 CC Query Match 55.3%; Score 1154; DB 1; Length 416;
 CC Best Local Similarity 56.6%; Fred. No. 2.3e-90;
 CC Matches 223; Conservative 65; Mismatches 94; Indels 12; Gaps 7;
 CC
 CC 10 LCCLALALALP-AQSCGPGVGRVRRARAKOLVLLYKOFVGPVPTLTLASGPAEGRVA 68
 CC 13 LCILSLTLTCGLACGGRG-YGKRHRPK-LTPLAYKQIIPNVAETILGASKEYEKIT 70
 CC 69 RSGSERFDVLPNTNPDIIFDEENSGADRLMTCKERKERNALALAVNMMPGVRLRTGEG 128
 CC 71 RNSERFELIPNPDIIIFDEENTNADRLMTKCKDKLNSLAVNMHPGVRLFTVEG 130
 CC 129 WDDGHAOSLYEGRALDITSDRDKRYGGLARLAVAGPFWYVESRHHVSVKA 188
 CC 131 WDEGHLLESLHVEGRAVDITSDRDKSKYGMISRLAVEAGFWMYVESKAHHCYSKA 190

CC 189 DNSLAVRAGCGEPGNATVRLMSGKRGLEHARGDVLADASGRVPPVLLFLDRDQ 248
 CC 191 ENSVAASKSGCGFPGSGTGLGDGTRKPIDKRGDRLVLADEKGNVLSDFIMEIDHPT 250
 CC 249 RASFLAVETEMPRKLLTLPMLVFAAGPAPAPDPAFVARRLRAGDSVLA--PGGD 306
 CC 251 TRQFTVIEFTEPFTLLTLTAALVEV--GNSSAASGATATFAFNVKPGDTLVWEDTCE 308
 CC 307 ALRPARVAVAREAVGFAPLPAHGTLLVNDVASCYAVLESNOVAHRAFAPLRLHAL 366
 CC 309 SLKSVYKRIYTEHESGAPVTAHGTIYDVYLACIYAVIENKHAHNAFAVRCLHL 368
 CC 367 GA-LDPGA---VQPTGMHYSRLLYRAELL 395
 CC 369 MTWLFPARESNVNFQDGIHWYSNMLFIQSWLL 402
 CC
 CC RESULT 13
 CC ID SHH_CYNPY STANDARD; PRT; 432 AA.
 CC AC 090385;
 CC DT 15-JUL-1999 (Rel. 38, Created)
 CC DT 15-JUL-1999 (Rel. 38, Last sequence update)
 CC DT 15-JUL-1999 (Rel. 38, Last annotation update)
 CC DE SONIC HEDGEHOG PROTEIN PRECURSOR (SHH).
 CC GN SHH.
 CC OS Cynops pyrrhogaster (Japanese common newt).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
 CC OC Batrachia; Caudata; Salamandridae; Salamandridae; Cynops.
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC TISSUE-EMBRYO;
 CC RA MEDLINE: 96136334.
 CC RX Takahashi T., Takahashi T.C., Inoue K., Ogawa M., Takeshima K.;
 CC RT "Activation of two Cynops genes, fork head and sonic hedgehog, in
 CC animal cap explants."
 CC RL Biochem. Biophys. Res. Commun. 218:395-401(1996).
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT AND METAMORPHOSIS. INVOLVED
 CC IN LIMB FORMATION, PATTERNING OF THE CENTRAL NERVOUS SYSTEM AND
 CC VENTRAL SOMITE DIFFERENTIATION. INDUCES ECTOPIC CEMENT GLAND
 CC FORMATION IN EMBRYOS. BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH
 CC FUNCTIONS IN ASSOCIATION WITH SMOOTHED (SMO), TO ACTIVATE THE
 CC TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH, PTC
 CC REPRESENTS THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- INDUCTION: ACTIVATED BY ACTIVIN, BASIC FIBROBLAST GROWTH FACTOR
 CC (bFGF) AND FORK HEAD.
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sid.ch).
 CC
 CC EMBL: D63339; BAA09657.1; -

PFAM; PF01079; Hint: 1.
 DR PFAM; PF01085; HR_Signal: 1.
 DR PRINTS; PR00632; SONICHHOG.
 KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 Signal.
 FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 27 432 SONIC HEDGEHOG PROTEIN.
 FT CHAIN 27 200 SONIC HEDGEHOG PROTEIN N-PRODUCT.
 FT CHAIN 201 432 SONIC HEDGEHOG PROTEIN C-PRODUCT.
 FT SITE 200 201 CLEAVAGE (AUTO-) (BY SIMILARITY).
 FT SITE 268 268 INVOLVED IN AUTO-CLEAVAGE (BY
 SIMILARITY).
 FT ACT_SITE 271 271 ESSENTIAL FOR AUTO-CLEAVAGE (BY
 SIMILARITY).
 FT BINDING 200 200 CHOLESTEROL (BY SIMILARITY).
 FT SEQUENCE 432 AA; 47847 MW; B455C7E746C8E5A8 CRC64;
 SO
 Query Match 54.2%; Score 1131.5; DB 1; Length 432;
 Best Local Similarity 55.5%; Pred. No. 1.9e-8;
 Matches 234; Conservative 57; Mismatches 96; Indels 35; Gaps 9;
 QY 1 MALLINLP---LCCIALALPAQSCGPGVRRYARKQVLPVLYKQVGPVPTL 57
 DB 4 MILLARVLLAGTIC--ALLVPSGLSCGPRGIGTKRF--KLTPLAYKQFTPNVPEKTL 59
 QY 58 GASGPAEGRVANGSERFNDLVPNYNDIIFKDENSGADRLMTERCKERNALAIAYNM 117
 DB 60 GASGVEGKITNSERFKELTPNYNDIIFKDEENTGADRLMTQCKDKRLNALAIAYNM 119
 QY 118 WGVGLRVTEGDEDEGHADSDIAYEGRALDITSDRONKKGILARLAVENGFPWYIE 177
 DB 120 WGVGLRVTEGDEDEGHAEESLHTEGRAVDITTDROSKTGMARLAAEGFPWYIE 179
 QY 178 SRNHVVSADKNSIAYVAGCGCPGNATVRLWSGERKGLREIHRGDWVYLAADSGRVPT 237
 DB 180 SKAHHCVCVKAENSVAAASGCGFPSSATVVALEQGVRIPIKDLRPDRVLAADGLKLYIS 239
 QY 238 PVLFLDLRLORASFAVETEMPRLKLLPWLHVA--ARGPAPADGFAPV----- 290
 DB 240 DELLFMDKEETRYKVFYVYIETS--REVRVLTAAHLFLFGQAHFGNDSGDFSVGSAGF 297
 QY 290 ---FARRRAGDSVAPG--GDALRPARVARAREAVGVPAPLPAHGLLVNDVLACY 344
 DB 298 RSMFASVVRAGRVRLVDREGRLAEATVERVLEATGAAPVPAHGTIVYDRLACY 357
 QY 345 AVLESHQWNAHRAFAPLRLHLAGAL-----LPGAVQPTGMWMSLYRLAE 393
 DB 358 AVIEHSHMAHNAFAPLRV--GLGILSFTSPQDYSSHPAPSGGVHMYSEILYRIGTW 415
 QY 394 LL 395
 DB 416 VL 417
 RESULT 14
 SHH_BRARE STANDARD; PRT; 418 AA.
 AC 092008: 013170: 013171:
 DT 15-JUL-1999 (Rel. 38, Last Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE SONIC HEDGEHOG PROTEIN PRECURSOR (SHH) (VHH-1).
 GN SHH OR VHH1.
 OS Brachydanio rerio (zebrafish) (zebra danio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
 CC Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
 CC Cyprinoidae; Cyprinidae; Rasbora; Danio.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-EMBRYO;
 RX MEDLINE; 94170375.
 RA Roelink H., Augsburg J., Heemskerk J., Korzh V., Norlin S.,

RA Ruiz I Altaba A., Tanabe Y., Placzek M., Edlund T., Jessell T.M.,
 RA Dodd J.;
 RT "Floor plate and motor neuron induction by vhh-1, a vertebrate
 RT homolog of hedgehog expressed by the notochord.";
 RL Cell 76:761-775(1994).
 RN [2]
 RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.
 RX MEDLINE; 96014264.
 RA Ekker S.C., Ungar A.R., Greenstein P., von Kessler D.P., Porter J.A.,
 RA Moon R.T., Beachy P.A.;
 RT "Patterning activities of vertebrate hedgehog proteins in the
 RT developing eye and brain.";
 RL Curr. Biol. 5:944-953(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 96083328.
 RA Fietz M.J., Concordet J.-P., Barbosa R., Johnson R., Krauss S.,
 RA McMahon A.P., Tabin C., Ingham P.W.;
 RT "The hedgehog gene family in Drosophila and vertebrate development.";
 RL Development Suppl. 43-51(1994).
 RN [4]
 RP SEQUENCE OF 30-92 AND 113-170 FROM N.A.
 RC TISSUE-MUSCLE;
 RX MEDLINE; 97075114.
 RA Zardoya R., Abouheif E., Meyer A.;
 RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species
 RT closely related to the zebrafish.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED BY THE
 CC NOTOCHORD THAT INDUCES SOMITE PATTERNING, DORSO-VENTRAL PATTERNING
 CC OF THE BRAIN AND EARLY PATTERNING OF THE DEVELOPING EYES. DISPLAYS
 CC FLOOR PLATE-INDUCING ACTIVITY. BINDS TO THE PATCHED (PTC)
 CC RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHED (SMO), TO
 CC ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH,
 CC PTC REPRESSSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE VENTRAL MIDLINE OF THE NEURAL
 CC TUBE AND BRAIN. ALSO FOUND IN THE NOTOCHORD AND IN DEVELOPING FIN
 CC BUD. IN THE DEVELOPING BRAIN, EXPRESSION OCCURS IN DOMAINS THAT
 CC INCLUDE A DISCRETE REGION IN THE FLOOR OF THE DIENCEPHALON.
 CC -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE IN THE INNER CELL LAYER OF
 CC THE EMBRYONIC SHIELD DURING GASTRULATION. BY 9.5 HR OF
 CC DEVELOPMENT, EXPRESSED IN A CONTINUOUS BAND THAT EXTENDS FROM THE
 CC TAIL TO THE HEAD, THE ANTERIOR BOUNDARY OF EXPRESSION BEING
 CC POSITIONED IN THE CENTER OF THE ANIMAL POLE ANTERIOR TO THE
 CC PRESUMPTIVE MIDBRAIN.
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; L27585; AAA20998.1; -
 DR EMBL; U30711; AAC59742.1; -
 DR EMBL; Z35669; CAA84738.1; -

EMBL: U51351; AAB38575.1; -;
 DR EMBL: U51370; AAB38593.1; -;
 DR ZFIN: ZDB-GENE-980526-166; SHH.
 DR PFAM: PF01079; Hntc. 1.
 DR PFAM: PF01085; HH_signal; 1.
 DR PRINTS: PRO0632; SONICHOG.
 KM Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 signal.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 418 SONIC HEDGEHOG PROTEIN.
 FT CHAIN 24 197 SONIC HEDGEHOG PROTEIN N-PRODUCT.
 FT CHAIN 198 418 SONIC HEDGEHOG PROTEIN C-PRODUCT.
 FT SITE 197 198 CLEAVAGE (AUTO-).
 FT SITE 243 243 INVOLVED IN CHOLESTEROL TRANSFER (BY
 SIMILARITY).
 FT SITE 267 267 INVOLVED IN AUTO-CLEAVAGE (BY
 SIMILARITY).
 FT ACT_SITE 270 270 ESSENTIAL FOR AUTO-CLEAVAGE (BY
 SIMILARITY).
 FT BINDING 197 197 CHOLESTEROL (BY SIMILARITY).
 FT SEQUENCE 418 AA; 46402 MW; CF000AFED2F5795 CRC64;
 SQ
 Query Match 53.68; Score 1120; DB 1; Length 418;
 Best Local Similarity 54.38; Pred. No. 1.7e-87;
 Matches 220; Conservative 66; Mismatches 107; Indels 12; Gaps 5;
 1 MALTLNLLPCLCAL-LALPAOSGPGPGVRRYARKOLVPLLYKQVGPVPTGGA 59
 1 MLRLTRVLVSLTLVSLVSGLAGCPGPG-YGRRHRPK-LTPLAKQFIVNAEKITGA 58
 60 SGPAERGVARGSEFRDLVPIYNDIIFKDENSGADRLMERKERVNALAVMMNP 119
 59 SGREGKITRSEFKELTPYNDIIFKDENTGADRLMTQROCKDLSIAIVMNP 118
 120 GYRLRVTEGMEDEHHAODSLHTEGRALDITSDRDKRYGLRLAVEAGFDWYYSR 179
 119 GYRLRVTEGMEDEHHAODSLHTEGRALDITSDRDKRYGLRLAVEAGFDWYYSR 178
 180 NHVAVYKADSLAVRAGGCFPGNATVRLMGERKGLBELRGDMVLAADSGRVPTPV 239
 179 AHICSVKAEVSAKSGCGFPGSALVSLQGGOKAVADLNPGRKYLAAASGNLVESDF 238
 240 LFLDRDLQRRASFAVETEMPRLTLTPHVLVFAARGAPADGFAVFAARLQADS 299
 239 IMFDRDSTTRVRYVLTETQEPVEKITLTAHLFLVDLNSDEDLHTMAAASSVRAQK 298
 300 VL-APGGDALPARYARAREAVGVPAPLTAHGTLLVNDVLAACYAVLASHQMAHRAFA 358
 299 VMVVDGSQLSVIVORLYTEORSEFAPVTAHGTIVDRILASCYAVIEDQGLAHLAFA 358
 359 PLRLHALGALL-----PGAVOPTGMHYSRLVRLABELL 395
 359 PARLYYVSFLFPONSSSRNATLQGGVHMYSLRLYKGTWLL 403
 RESULT 15
 SHH_XENLA STANDARD; PRT; 444 AA.
 AC Q93000; Q91894;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-FEB-2000 (Rel. 39, Last annotation update)
 DE SONIC HEDGEHOG PROTEIN PRECURSOR (X-SHH) (YHH-1).
 GN SHH.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
 OC Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae;
 OC Xenopus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-INTERESTING;
 RX MEDLINE; 95357169.

RA Stelow M.A., Shi Y.-B.;
 RT "Xenopus sonic hedgehog as a potential morphogen during embryogenesis
 RT and thyroid hormone-dependent metamorphosis."
 RL Nucleic Acids Res. 23:2555-2562(1995).
 RN [2]
 { RP SEQUENCE FROM N.A.
 RC TISSUE-EMBRYO;
 RX MEDLINE; 95401852.
 RA Erker S.C., McGrew L.L., Lai C.-J., Lee J.J., von Kessler D.P.,
 RA Moon R.T., Beachy P.A.;
 RT "Distinct expression and shared activities of members of the hedgehog
 RT gene family of Xenopus laevis."
 RL Development 121:2357-2347(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE-NOTOCHORD;
 RX MEDLINE; 96028338.
 RA Ruiz T. Alcoba A., Jessell T.M., Roelink H.;
 RT "Restrictions to floor plate induction by hedgehog and winged-helix
 RT genes in the neural tube of frog embryos."
 RL Mol. Cell. Neurosci. 6:106-121(1995).
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT AND METAMORPHOSIS. INVOLVED
 CC IN LIMB FORMATION. PATTERNING OF THE CENTRAL NERVOUS SYSTEM AND
 CC VENTRAL SOMITE DIFFERENTIATION. INDUCES ECTOPIC CEMENT GLAND
 CC FORMATION IN EMBRYOS. BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH
 CC FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE
 CC TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH, PTC
 CC REPRESENTS THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: SPONGY EXPRESSED IN NOTOCHORD AND NEURAL
 CC FLOOR PLATE DURING EMBRYOGENESIS. IN TADPOLE, HIGH EXPRESSION IS
 CC OBSERVED IN PANGRAS/STOMACH, MODERATE EXPRESSION IN TAIL, AND LOW
 CC EXPRESSION IN INTESTINE, BRAIN, AND HIND LIMB.
 CC -1- DEVELOPMENTAL STAGE: FIRST DETECTED AT THE NEURULA (STAGES 16-17).
 CC FIRST PEAK OF EXPRESSION AROUND TADPOLE HATCHING (STAGES 33-40).
 CC HIGH EXPRESSION OBSERVED IN INTESTINE AT THE CLIMAX OF
 CC MORPHOGENESIS (STAGES 60-62) WHEN INTESTINE EPITHELIAL UNDERGOES
 CC MORPHOGENESIS.
 CC -1- INDUCTION: BY THYROID HORMONE.
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL collaboration.
 CC The European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L39213; AAC4227.1; -;
 DR EMBL: U26314; AAA85162.1; -;
 DR EMBL: L35248; AAA4981.1; -;
 DR PFAM: PF01079; Hntc. 1.
 DR PFAM: PF01085; HH_signal; 1.
 DR PRINTS: PRO0632; SONICHOG.
 KM Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 FT SIGNAL
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 444 SONIC HEDGEHOG PROTEIN.

```
FT CHAIN 25 198 SONIC HEDGEHOG PROTEIN N-PRODUCT.
FT CHAIN 199 444 SONIC HEDGEHOG PROTEIN C-PRODUCT.
FT SITE 198 199 CLEAVAGE (AUTO-) (BY SIMILARITY).
FT SITE 266 266 INVOLVED IN AUTO-CLEAVAGE (BY
FT ACT_SITE 269 269 ESSENTIAL FOR AUTO-CLEAVAGE (BY
FT DOMAIN 386 409 SIMILARITY).
FT REPEAT 386 393 3 X 8 AA TANDEM REPEATS OF Q-V-D-L-Q-S-H-
FT REPEAT 394 401 1.
FT REPEAT 403 409 2.
FT BINDING 198 198 3.
FT CONFLICT 5 9 CHOLESTEROL (BY SIMILARITY).
FT CONFLICT 302 319 TOSL -> NSNLCW (IN REF. 3).
FT CONFLICT 302 319 DPKMTLKAVKEKYDLE -> ESQDHLEGRGKWRRLILR
FT CONFLICT 432 432 (IN REF. 3).
FT CONFLICT 432 432 N -> S (IN REF. 3).
SQ SEQUENCE 444 AA; 49453 MW; 73B4E4932FA2EFF2 CRC64;
```

Query Match 52.3%; Score 1092.5; DB 1; Length 444;
Best Local Similarity 50.1%; Pred. No. 4e-85;
Matches 213; Conservative 71; Mismatches 100; Indels 41; Gaps 7;

```
QY 7 LPLCLALALPAQSCGPGRGVGRRRYARKQVLPILYKQFVGPVPTLQASGPAEGR 66
DB 10 LLSFIC-TLVTPGLACGPGRG-IGKRHRPK-LTPAYKQFIPNVAEKTIGASGRYEGK 66
QY 67 VARGSERFDLVNPNPDIFKDEENSGADRLMTERCKEKNALATAVNMMPGVRLRYT 126
DB 67 ITNSOCFKELITPNYNDIFKDEESTGADRLMTERCKEKNALATAVNMMPGVRLRYT 126
QY 127 EGVDEGHHNAQDSLHYEGRALDITSDRDKYGLRLAVEAGFDVYYESRNHVSV 186
DB 127 EGVDEGHHLESLHYEGRAVDITSDRDKYGLRLAVEAGFDVYYESKAHHCYV 186
QY 187 KAPNSLAVRAGCGFPGNATYRLMSGERKGLRELRGDMVLAADSGRVVPTVLLFLDRD 246
DB 187 KAPNSVAAKSGCGFPGARVYVEFGTKAVKDLRPGDRVLSDPQGNLLYSDFLMFIDQE 246
QY 247 LQRRASFVAVETWPPRKLLLTWHLVFAARGPAPAGDFAPFARRLRAGDSVLA--PG 304
DB 247 RQVKLFLYIETS--QRKILTAHLLFVAQTVNGTRSFKSYFASNIQPGDLITYADPR 304
QY 305 GDALRPARAVAREAVGVFAPRLTAGTLVNDVLASCYAVLESHQWARRAPRLRLH 364
DB 305 TMTLKAVKVEKYDLEDTGAYAPLTAAGTVIDQVLASCYAVIEHTWALAPRLRFGM 364
QY 365 ALGALL-----PGAVQP-----TGMWTSRLLYRL 390
DB 365 SLSSYIYPRDSSPPSGIQRHQVQLQSHQVQLQSHQVQLQSHQLEGIHWYSQLLYQI 424
QY 391 AEELL 395
DB 425 GTWLL 429
```

Search completed: June 5, 2000, 08:19:55
Job time: 159 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2000, 08:19:30 ; Search time 26.41 Seconds
(without alignments)
1039,618 Million cell updates/sec

Title: US-08-900-220-17
Perfect score: 2088
Sequence: 1 MALTNLPLCLALLALPA.....PTGMWYSRLYLAEELG 396

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 225878 seqs, 6934122 residues
Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :
1: SPTREMBL.12.*
2: SP_Archea.*
3: SP_Bacteria.*
4: SP_Fungi.*
5: SP_Human.*
6: SP_Invertebrate.*
7: SP_Mammal.*
8: SP_Mhc.*
9: SP_Organelle.*
10: SP_Phage.*
11: SP_Plant.*
12: SP_Protoct.*
13: SP_Virus.*
14: SP_Vertebrate.*
15: SP_Unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1144.5	54.8	406	13	057567 notophthalm
2	1126	53.9	414	13	09W709 parolichthy
3	1107.5	53.0	434	13	057404 pleurodeles
4	990	47.4	442	13	073803 fugu rubrip
5	919	44.0	415	5	017499 branchiost
6	781.5	37.4	410	5	061676 lytechinus
7	664	31.8	129	11	09W0P6 ratcus norv
8	624	29.9	177	11	09W29 ratcus norv
9	618	29.6	139	6	09XSi6 bos taurus
10	606	29.0	150	13	09YGV7 ambystoma m
11	599	28.7	138	13	09W6C1 eleutheroda
12	585	28.0	185	5	096699 junonia coe
13	431	20.6	119	13	042128 oryzias lat
14	375	18.0	88	13	09YGV3 brachydanio
15	352	16.9	80	13	042441 oryzias lat
16	273.5	13.1	137	13	042234 coturnix co
17	209	10.0	868	5	09XV14 caenorhabdi
18	194.5	9.3	790	5	022872 caenorhabdi
19	184.5	9.3	1226	5	021835 caenorhabdi
20	193.5	9.3	1021	5	09XUV2 caenorhabdi

21	186.5	8.9	1207	5	021535 caenorhabdi
22	182	8.7	481	5	045992 caenorhabdi
23	179.5	8.6	557	5	094129 caenorhabdi
24	163	7.8	485	5	094128 caenorhabdi
25	161.5	7.7	550	5	094130 caenorhabdi
26	161.5	7.7	629	5	045273 caenorhabdi
27	150.5	7.2	615	5	P91573 caenorhabdi
28	135	6.5	205	5	023193 caenorhabdi
29	115.5	5.5	3670	2	0924X5 streptomyce
30	114.5	5.5	687	2	P71196 escherichia
31	106.5	5.1	481	5	094410 caenorhabdi
32	103.5	5.0	3391	12	092834 dengue viru
33	103.5	5.0	3391	12	092835 dengue viru
34	103	4.9	672	2	09X856 streptomyce
35	102.5	4.9	3391	12	01875 dengue viru
36	102.5	4.9	3391	12	009234 unidentified
37	102.5	4.9	3391	12	092752 dengue viru
38	102.5	4.9	3391	12	092753 dengue viru
39	102.5	4.9	3391	12	09W1Z8 dengue viru
40	102.5	4.9	3391	12	09W1Z7 dengue viru
41	102.5	4.9	3391	12	09W1Z4 dengue viru
42	102.5	4.9	3391	12	09WDA7 dengue viru
43	102.5	4.9	3391	12	09WDA6 dengue viru
44	102.5	4.9	3391	12	09WDA5 dengue viru
45	102.5	4.9	3391	12	09WDA4 dengue viru

ALIGNMENTS

RESULT	1	PRELIMINARY;	PRT;	406 AA.
ID	057567			
AC	057567			
DT	01-JUN-1998 (TREMBLrel. 06, Created)			
DT	01-JUN-1998 (TREMBLrel. 06, Last sequence update)			
DT	01-NOV-1999 (TREMBLrel. 12, Last annotation update)			
DE	HEDGEHOG SEGMENT POLARITY HOMOLOG.			
OS	Notophthalmus viridescens (Eastern newt) (Triturus viridescens).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;			
CC	Batrachia; Caudata; Salamandridae; Salamandridae; Notophthalmus.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	STARK D.R., GATES P.B., BROCKES J.P., FERRETTI P.,			
RL	Dev. Dyn. 0:0-0(1998).			
DR	EMBL; AF047466; AAC03108.1; -.			
DR	HSSP; Q62226; YVH.			
DR	PFAM; PF01085; HH_signal; 1.			
DR	PFAM; PF01079; HInt; 1.			
DR	PRINTS; PR00632; SONICHOG.			
DR	SEQUENCE 406 AA; 45072 MW; 9D0FFA76 CRC32;			

Query Match	54.8%	Score 1144.5;	DB 13;	Length 406;
Best Local Similarity	58.2%	Pred. No. 2.1e-81;		
Matches 227;	Conservative 49;	Mismatches 107;	Indels 7;	Gaps 3;
QY	7	LPLCLALLALP-NQSCGPGRGVRRYARKQVLPLTKQFVGVPERTIGASPAEG	65	
DB	8	LVAACALLLGPGALCGPGR-VIGRRPRPLIPSTYQFPLHYEKTIGASGRYEG	65	
QY	66	RVANGSERFDLPVNPYNDIFKDENSGADRLMTERCKEYVNLALAVMMAMGVRRLRV	125	
DB	66	KIANSEFEKLTNPYNDIFKDENTGADRLMTERCKEYVNLALAVMMAMGVRRLRV	125	
QY	126	TEGWDEGHHADSLHYEGRALDITTSDDRNKXGLARLAVEAGFWMYVESNHHVHS	185	
DB	126	TEGWDEGHHADSLHYEGRALDITTSDDRNKXGLARLAVEAGFWMYVESNHHVHS	185	
QY	186	VKAANSIAVNRGCGFPFGATYRLMSGERKGLREHGRDWTYLAASGRVPTPLTLDR	245	
DB	186	VKSHSAAKTGGCFPAALATLESGERIPADLEPGHRYLCMDGGRRTYSDFLTFIDR	245	

QY 246 DLORRASFAVATEMPPRKLLTPMHLVFAARGAPAPGDAFVAPARLRLAGSDVLAFCG 305
DB 246 DSRVAEEFYVERDRPPRLALTAHLFTADNFTVPLDFTSFVFSHPQGYILLTEGV 305
QY 306 DALRPARVARAEEBAVAPLTAAGTLLVNDVLAACYAVLASHOMAHAFAPRLRLHA 365
DB 306 LGLOPARVAVSVTQTDGSAAPLTLSHOTLLVDVYVSCFAVOKHOLAQAFAPRLRLYHS 365
QY 366 LGALLPEGAVOPTGMHWYSLLYRLAEELL 395
DB 366 VGR-----PETQPEGMHWYSLLYRLKGVLL 391

RESULT 2

QY 09W709 PRELIMINARY: PRT: 414 AA.
AC 09W709;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE SONIC HEDGEHOG.
GN SHH.
OS Paralicthys olivaceus (Flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
OC Pleuronectiformes; Pleuronectoidel; Bothidae; Paralicthys.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 99238226.
RA SUZUKI T., ICHIRO O., KUROKAWA T.;
RT "Retinoic acid given at late embryonic stage depresses sonic hedgehog
RT and Hoxd-4 expression in the pharyngeal arch and induces skeletal
RT malformation in flounder (Paralicthys olivaceus) embryos.";
RL Dev. Growth Differ. 41:143-152(1999).
RL EMBL: AB029748; BAA82360.1;
SQ SEQUENCE 414 AA: 45945 MW: 51FBI2EE CRC32;

Query Match

Best Local Similarity 53.9%; Score 1126; DB 13; Length 414;
Matches 218; Conservative 71; Mismatches 96; Indels 10; Gaps 6;

QY 7 LRLPLCCALLALPAQSGGPGRGVRRRRYARKQVLPVLYKQFVGPVPTLGAAGPAEGR 66
DB 9 LAGVITCLSLVS-SGMCQGGPGRG-YGRRRHPKK-LTPLAYKQFIPNVAEKTGLASGRYEGK 65
QY 67 VARGSERFDLVENYNDIIFKDENSGADRLMTERCKEKNALATAVNMMPGVRLRY 126
DB 66 ITRNSERFKELTPNYNDIIFKDEMTGADRLMTQRCOKDLNSLAISVMQWPGVRLRY 125
QY 127 EGDGDEGHAQDSLHYEGRALDITTSDDRNKYGGLARLAVEAGFDMVYYESNNHVS 186
DB 126 EGDGDEGHAQDSLHYEGRALDITTSDDRNKYGGLARLAVEAGFDMVYYESNNHVS 185
QY 187 KADNSLAVRAGGCGFPGNATVRLWSGKGLRELHBDGMDVLAASGRVPTPLFLDRL 246
DB 186 KADNSLAVRAGGCGFPGNATVRLWSGKGLRELHBDGMDVLAASGRVPTPLFLDRL 245
QY 247 LORRASFAVATEMPPRKLLTPMHLVFAARGAP--APGDAFVAPARLRLAGSDVLAFCG 304
DB 246 STRRLFYVLETD-SGOKITLLTAHLTFVGHSHSTERAHNSMAVYASQVRFQGYTFVLD 304
QY 305 GDALRPARVARAEEBAVAPLTAAGTLLVNDVLAACYAVLASHOMAHAFAPRLRLH 364
DB 305 AERLQAVTVARITQHEGSAFAVTAQGVVDOVLASCYAVIADHBLAHMALAPVRLAH 364
QY 365 ALGALL-----PGAVOPTGMHWYSLLYRLAEELL 395
DB 365 WVSLLFSSQPAQAKDGVHWYSKILYOLGTWLL 399

RESULT 3
057404

ID 057404 PRELIMINARY: PRT: 434 AA.
AC 057404;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE SONIC HEDGEHOG-RELATED PROTEIN.
GN PW-SHH.
OS Pleurodeles waltlil (Iberian ribbed newt).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC Batrachia; Caudata; Salamandroidae; Salamandridae; Pleurodeles.
RN [1]
RP SEQUENCE FROM N.A.
RA CAUBIT X., NICOLAS S., LE PARCO Y.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF005332; AAB94412.1;
DR HSSP: Q62226; 1VH.
DR PFAM: PF01085; HH.signal; 1.
DR PFAM: PF01079; Hint; 1.
DR PRINTS: PR00632; SONICHHOG.
SQ SEQUENCE 434 AA: 48421 MW: AFFFD0F3 CRC32;

Query Match

Best Local Similarity 53.0%; Score 1107.5; DB 13; Length 434;
Matches 225; Conservative 59; Mismatches 91; Indels 43; Gaps 8;

QY 5 TNLPLCCALLALPAQSGGPGRGVRRRRYARKQVLPVLYKQFVGPVPTLGAAGPAE 64
DB 18 TLLVPL-----GLGCGPGRG-IGQRRRPQK-LTPLAYKQFIPNVAEKTGLASGRYE 66
QY 65 GVARASERFDLVENYNDIIFKDENSGADRLMTERCKEKNALATAVNMMPGVRLRY 124
DB 67 VITRSERFKELTPNYNDIIFKDEMTGADRLMTQRCOKDLNSLAISVMQWPGVRLRY 126
QY 125 VEGMDDEGHAQDSLHYEGRALDITTSDDRNKYGGLARLAVEAGFDMVYYESNNHVS 184
DB 127 VEGMDDEGHAQDSLHYEGRALDITTSDDRNKYGGLARLAVEAGFDMVYYESNNHVS 186
QY 185 SVKADNSLAVRAGGCGFPGNATVRLWSGKGLRELHBDGMDVLAASGRVPTPLFLD 244
DB 187 SVKADNSLAVRAGGCGFPGNATVRLWSGKGLRELHBDGMDVLAASGRVPTPLFLD 246
QY 245 ROLORRASFAVATEMPPRKLLTPMHLVFAARGAPAGD-----FAPVF 290
DB 247 EERLAKFYVLETD-SGOKITLLTAHLTFVGHSHSTERAHNSMAVYASQVRFQGYTFVLD 304
QY 291 AERLQAVTVARITQHEGSAFAVTAQGVVDOVLASCYAVIADHBLAHMALAPVRLAH 363
DB 304 AERLQAVTVARITQHEGSAFAVTAQGVVDOVLASCYAVIADHBLAHMALAPVRLAH 363
QY 349 SHOMAHAFAPRLRLHAGAL-----LPGAVOPTGMHWYSLLYRLAEELL 395
DB 364 SHOMAHAFAPRLRLHAGAL-----LPGAVOPTGMHWYSLLYRLAEELL 395
RESULT 4
073803
ID 073803 PRELIMINARY: PRT: 442 AA.
AC 073803;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE FUGU HEDGEHOG.
GN FHH.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
OC Tetraodontiformes; Tetraodontoidel; Tetraodontidae; Fugu.
RN [1]
RP SEQUENCE FROM N.A.
RA GELINER K., BRENNER S.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF056116; AAC34384.1;

QY 307 ALPRPRVARAREEAVGFAPLLTNGHSTLVNDVLAACVALESQOMHRAAPR-----L 3662

Db 309 GVRPRKVVSVTTLRGHTAVAPVTPQGSLYIDGVAVSSAVARDEMAHNASAPIRWTTYI 3688

QY 363 LHALGALLPGGANVPTGTHMHTSKLLYLRAEELL 395

Db 369 SHMUG-LITDITLGGEOQRVHMTQGLYKIGKTYM 400

RESULT	7
Q9WUP6	
ID	Q9WUP6 PRELIMINARY; PRT; 129 AA.
AC	Q9WUP6;
DT	01-NOV-1999 (TrEMBLrel. 12, Created)
DT	01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT	01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE	DESERT HEDGEHOG PROTEIN (FRAGMENT).
CN	DHH.
OS	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia
OC	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=SPRACUE-DAMELEY;
RA	GARGES P.L., MEYER R.A. JR., BROWN C.A., PRICE D.K.:
RT	"Desert hedgehog in the rat."
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF148226; AACD3127.1; -
FT	NON_TER 1
FT	NON_TER 1
SO	SEQUENCE 129 AA; 14578 MW; CB8B2D40 CRC32;

RESULT	8		
09WV29			
ID	09WV29	PRELIMINARY;	PRT: 177 AA.
AC	09WV29;		
DT	01-NOV-1999 (TREMBLrel. 12, Created)		
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)		
DT	01-NOV-1999 (TREMBLrel. 12, Last annotation update)		
DE	INDIAN HEDGEHOG PROTEIN (FRAGMENT).		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;		
NC	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-SPRAGUE-DAWLEY; TISSUE-KIDNEY;		
RA	GARGES P. L., MEYER R.A. JR., BROWN C.A., PRICE D.K.;		
RT	"Indian hedgehog in rat."		
RL	Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.		
DR	EMBL: AF162914; A4D45372.1; -.		
FT	NON_TER	1	1
FT	NON_TER	177	177
SQ	SEQUENCE	177 AA;	19739 MW; 24CF1044 CRC32;

Query Match	29.9%	Score 624	DB 11	Length 177
Best Local Similarity	65.5%	Pred. No. 2	1e-41	
Matches 116	Conservative 25	Mismatches 36	Indels 0	Gaps 0
QY	18	INTERCKERYNMLATAVAMNMPGVRLRYTEGDEWDEGHHADBSLHAYEGRALDITTSDDRN	157	
DB	1	LTGQCKDRNLNSLAISVNWQMPGVRLRYTEGDEGDEGHHSESLHAYEGRALDITTSDDRN	60	
QY	158	KYGLLARLAVLEGFWYTYESENHYVSKADNSLAVAGCGFCGNATVRLMSGERKCLR	217	
DB	61	KYGLLARLAVLEGFWYTYESENHAYCSKSHSAAATGCGFPAGAVVHETGERVALS	120	
QY	218	ELHRGDVLADASGRVYPTVLTLDLDDLORBSAFVAVETEMPRKLLTPMWHVE	274	
DB	121	AKVPDRVLAMGEDGNPTFSVLTITLDEDPNRLAFQVLETODPPRLATLPAHLLE	177	

RESULT	9			
09XS16				
ID	09XS16	PRELIMINARY;	PR1;	139 AA.
AC	09XS16:			
DT	01-NOV-1999	(TIREMBLrel. 12, Created)		
DT	01-NOV-1999	(TIREMBLrel. 12, Last sequence update)		
DT	01-NOV-1999	(TIREMBLrel. 12, Last annotation update)		
DE	SONIC HEDGEHOG (FRAGMENT).			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
OC	Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae;			
OC	Bovinae; Bos.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-TOOTH GERM;			
RA	KOYAMA E., IMAOTO M., OHMORI T., KURIKU K., WU C., OOKURA T.,			
RA	BAKHIR M.N., TOCKER T., PACIFICI M.;			
RT	"Development of Stratum Intermedium and its Role as a Sonic Hedgehog-			
RT	Signaling Structure During Odontogenesis."			
RL	Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF144100; AAD33926.1; -			
FT	NON_TER	1		
FT	NON_TER	139		
SO	SEQUENCE	139 AA;	15961 MW;	32DF025 CRC32;

RESULT	ID	Q9YGV7	PRELIMINARY:	PRF:	150 AA.
AC	Q9YGV7	01-MAY-1999	(TRIMBLrel. 10, Created)		
DT	01-MAY-1999	(TRIMBLrel. 10, Last sequence update)			
DT	01-NOV-1999	(TRIMBLrel. 12, Last annotation update)			
DE	SONIC HEADSHOG (FRAGMENT)				
OS	Ambystoma mexicanum (Axolotl).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Amphibia;				
CC	Batrachia; Caudata; Salamandroidea; Ambystomatidae; Ambystoma.				
RN	[1]				

THIS PAGE BLANK (usPTO)

THIS PAGE BLANK

Db	61	GPAGRGVYRSGSERRRDLVPPNNPILIFKDENSGADMLMERCKERNALALIVAMNMPG	120
QY	121	VLRLVTEGMDGDEGHHADDSLHYEGSRALDITTSDDRNNKYGLLARLAVEAGFDWYYESRN	180
Db	121	VLRLTEGMDGDEGHHAAQDSLHYEGSRALDITTSDDRNNKYGLLARLAVEAGFDWYYESRN	180
QY	181	HHVHSVKADNSLAIRAGCGFCGNNVTVLMSGKGLRELRGDMVYLAADSGRVVPTPYL	240
Db	181	HHHSVKADNSLAIRAGCGFCGNNVTVRLRSERKGLRELRGDMVYLAADAGRVVPTPYL	240
QY	241	LELDRLDQRRASFVAVEREMPPRKLTLTPMHLVFAARGPAPAPGDFEAPVARRLLRAGDSV	300
Db	241	LELDRLDQRRASFVAVEREMPRKLLTLTPMHLVFAARGPAPAPGDFEAPVARRLLRAGDSV	300
QY	301	LAPGCDALRPARVARVAREEAVGVFAPLTAHGTLLVNDVTLASCYAVLESHQMAHRAAPL	360
Db	301	LAPGCDALQPARVARVAREEAVGVFAFAPLTAHGTLLVNDVTLASCYAVLESHQMAHRAAPL	360
QY	361	RLTHALGLLPGSAGVOPGTGMWYSRLLYRLAEELG	396
Db	361	RLTHALGLLPGSAGVOPGTGMWYSRLLYRLAEELMG	396

```

RESULT 2
US-08-356-060A-9
Sequence 9, Application US/08356060A
Patent No. 5844079
GENERAL INFORMATION:
APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tabin, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
TITLE OF INVENTION: Proteins and Uses Related Thereto
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,060A
FILING DATE: 14-DEC-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/176,427
FILING DATE: 30-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMT-006CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 396 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-356-060A-9

```

Query Match	96.2%	Score 2008	DB 2	Length 396
Best Local Similarity	96.5%	Pred. No. 2.5e-214		
Matches 382	Conservative	6	Mismatches 8	Indels 0
			Gaps	0

Qy	1	MAATNNLLPCLCLALLLLPAGSCGGPGVGRRRYARKOLVPLLYIKQFYVGEVERTIGAS	60
Dp	1	MAIPASILLPCLCLALLLLASOCSGPGKPGRRRYVKOLVPLLYIKQFYVSMERTIGAS	60
Qy	61	GPAEGRVASGSEFRDLYPNPNPDIIIFKDEBNSGADRLMTERCKERNALAIAMNMMPG	120
Dp	61	GPAEGRVATRGSEFRRDLYPNPNPDIIIFKDEBNSGADRLMTERCKERNALAIAMNMMPG	120
Qy	121	VRLRVTEGWDGDEHHADDSLHYESRALDITTSDDRNNKYGLLARLAVEAGEFDWYYESRN	180
Dp	121	VRLRVTEGWDGDEHHADDSLHYESRALDITTSDDRNNKYGLLARLAVEAGEFDWYYESRN	180
Qy	181	HVHVSVKADNSLAVRAGCGPFGNATVYLMSEGRGELNEIRHGMVYLAADASGRVPPVYL	240
Dp	181	HHVSVKADNSLAVRAGCGPFGNATVYLRSGERKGLNEIRHGMVYLAADAGRVPPVYL	240
Qy	241	LELDRLDORRASFYVAVETEMPPRKLITLPMHLYFAARGPAPAPGDFAPVEARRLRACDSY	300
Dp	241	LELDRLDORRASFYVAVEDEPRRKLITLPMHLYFAARGPAPAPGDFAPVARRLRACDSY	300
Qy	301	LAPGCDALRPARVARVAREEAVGYFAPLTAHGTLLVNDVTLASCYAVLSEHQMAHRAAPL	360
Dp	301	LAPGCDALOPARVARVAREEAVGYFAEPLTAHGTLLVNDVTLASCYAVLSEHQMAHRAAPL	360
Qy	361	RLHAHLGALLPGSAVOPTGMWYSRLLYRLAAEELG 396	
Dp	361	RLHAHLGALLPGSAVOPTGMWYSRLLYRLAAEELMG 396	

RESULT 3
 US-08-116-427b-8
 Sequence 8, Application US/08176427B
 Patent No. 5789543
 GENERAL INFORMATION:
 APPLICANT: Ingham, Phillip W.
 APPLICANT: McMahon, Andrew P.
 APPLICANT: Tabin, Clifford J.
 TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
 TITLE OF INVENTION: Proteins and Uses Related Thereto
 NUMBER OF SEQUENCES: 33
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: LAHIVE & COCKFIELD
 STREET: 60 State Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: ASCII(text)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/176,427B
 FILING DATE: 30-DEC-1993
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Vincent, Matthew P.
 REGISTRATION NUMBER: 36,709
 REFERENCE/DOCKET NUMBER: HMI-006
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 227-7400
 TELEFAX: (617) 227-5941
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 437 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-116-427b-8

Query Match 58.3%; Score 1217; DB 1; Length 437;
 Best Local Similarity 58.7%; Pred. No. 1.6e-126;
 Matches 249; Conservative 53; Mismatches 82; Indels 40; Gaps 12;

```

QY 7 LPLPCCALTA-----LPASCGRGPGVGRRRARQVLPVLYKQVPGVPERLTASG 61
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 4 LLARCFVLIASSLVPCGACGPGRG-FGKRHPKR-LIPLAVKQFIIPVNAEKTIGASG 61
QY 62 PAEGRVARGSERFDLPVNPYNDIIFKDENSGADRLMTERCKERVNALAIVANNMPPGV 121
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 62 RYEGKITRNSRFKELPNNYNDIIFKDENTGADRLMTCCKKLNALASVNNQMPGV 121
QY 122 RLRTBEGWDEGHHADSLHYEGRALDITTSDRKNKYGILLARLAVAGFDWVYYESRNH 181
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 122 RLRTBEGWDEGHHSESLHYEGRAVDITTSDRSKYGMARLAVAGFDWVYYESKAH 181
QY 182 VHVSVKADNSLAVAGCGCPGNATVRLMSGKGLRELRHGDWVLADASGRVVPVTL 241
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 182 IHCSVKAENSVAAKSGCGFPGSATVHLEOGSTKLVKDLRPDRVLAADDOGRLLYSDFLT 241
QY 242 FLDRDLORASFAVETEMPRLKLLTPWHLVFAA---RGAPAPGDFAFVFAARRLRAG 297
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 242 FLDRDEGAKKVFYIETLEPRERLLTAHLLFVAPHNDSGPTGP---SALFASRVAPG 298
QY 298 DSVLA---PGGD-ALRAVARVA-REBAGVFAPLTAHGLLVNDVLAICYAVLESQHW 352
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 299 QRVYVAERGGDRLLPAVAHVTLRBEAGAYAPLTAHGLILNRYLASCYAVIEESHW 358
QY 353 AHRAFAPLRLHA-LGALLP-----GGAV-----OPT-GMHYSRLLYLA 391
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 359 AHRAFAPLRLHALLALAPARTDGGGGSIPAAOSATEARGAAPTACIHWYLSOLYHIG 418
QY 392 EELL 395
  ||
Db 419 TWLL 422

```

RESULT 4

```

US-08-356-060A-11
; Sequence 11, Application US/08356060A
; Patent No. 5844079
; GENERAL INFORMATION:
; APPLICANT: Ingham, Phillip W.
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Tablin, Clifford J.
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,060A
; FILING DATE: 14-DEC-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/176,427
; FILING DATE: 30-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMI-006CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400

```

TELEFAX: (617) 227-5941
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 437 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-356-060A-11.

Query Match 58.3%; Score 1217; DB 2; Length 437;
 Best Local Similarity 58.7%; Pred. No. 1.6e-126;
 Matches 249; Conservative 53; Mismatches 82; Indels 40; Gaps 12;

```

QY 7 LPLPCCALTA-----LPASCGRGPGVGRRRARQVLPVLYKQVPGVPERLTASG 61
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 4 LLARCFVLIASSLVPCGACGPGRG-FGKRHPKR-LIPLAVKQFIIPVNAEKTIGASG 61
QY 62 PAEGRVARGSERFDLPVNPYNDIIFKDENSGADRLMTERCKERVNALAIVANNMPPGV 121
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 62 RYEGKITRNSRFKELPNNYNDIIFKDENTGADRLMTCCKKLNALASVNNQMPGV 121
QY 122 RLRTBEGWDEGHHADSLHYEGRALDITTSDRKNKYGILLARLAVAGFDWVYYESRNH 181
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 122 RLRTBEGWDEGHHSESLHYEGRAVDITTSDRSKYGMARLAVAGFDWVYYESKAH 181
QY 182 VHVSVKADNSLAVAGCGCPGNATVRLMSGKGLRELRHGDWVLADASGRVVPVTL 241
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 182 IHCSVKAENSVAAKSGCGFPGSATVHLEOGSTKLVKDLRPDRVLAADDOGRLLYSDFLT 241
QY 242 FLDRDLORASFAVETEMPRLKLLTPWHLVFAA---RGAPAPGDFAFVFAARRLRAG 297
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 242 FLDRDEGAKKVFYIETLEPRERLLTAHLLFVAPHNDSGPTGP---SALFASRVAPG 298
QY 298 DSVLA---PGGD-ALRAVARVA-REBAGVFAPLTAHGLLVNDVLAICYAVLESQHW 352
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 299 QRVYVAERGGDRLLPAVAHVTLRBEAGAYAPLTAHGLILNRYLASCYAVIEESHW 358
QY 353 AHRAFAPLRLHA-LGALLP-----GGAV-----OPT-GMHYSRLLYLA 391
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 359 AHRAFAPLRLHALLALAPARTDGGGGSIPAAOSATEARGAAPTACIHWYLSOLYHIG 418
QY 392 EELL 395
  ||
Db 419 TWLL 422

```

RESULT 5

```

PCT-US95-15463-20
; Sequence 20, Application PC/TUS9515463
; GENERAL INFORMATION:
; APPLICANT: The Johns Hopkins University School of Medicine
; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/15463
; FILING DATE: 01-DEC-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347

```

REFERENCE/DOCKET NUMBER: 07265/080W01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-15463-20

Query Match 58.1%; Score 1214; DB 4; Length 437;
Best Local Similarity 58.5%; Pred. No. 3.5e-126;
Matches 248; Conservative 54; Mismatches 82; Indels 40; Gaps 12;

QY 7 LPLCLALLA-----LPAQSCGPGGPGVGRRRYARKOLVPLLYKQFVGYVPTIGASG 61
DB 4 LARCELYLASSLLVCPGLACGPRG-FGKRHPKK-LTPLAYKQFIPVAEKTIGASG 61
QY 62 PAGRVARGSERFDLVPNYNPDIIFKDEENSGADRLMTERCKERYNALAIAMNMPGV 121
DB 62 RYGRKTRNSERKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAIAMNMPGV 121
QY 122 RLRTGEMWEDGHHAQDSLYEGRALDITTSRDNRKYGILLARLAVEGFDWVYESRNH 181
DB 122 KLRTVEGWEDGHSESLHYEGRAVDITTSRDNRKYGILLARLAVEGFDWVYESRNH 181
QY 182 VHVSVKADNSLAVRAGCGPGNATVRLMSGERKGLRELHNGDMVLAADASGRVPTPYLL 241
DB 182 IHCSVAENSVAAGSGCGPGSATVHLEGGTKLVKDLRPGDVLADDOGRLLYSDFLT 241
QY 242 FLDRDLQRRASFVAVETEMPRLKLLTPMHLVFAA---RGPAAPGDFAPVAFARLRAG 297
DB 242 FLDRDGAQKRVYVETLEPRERLLTAHLFLVAPHNDGPTPGP---SALFASRVRG 298
QY 298 DSVLA---PGGD-ALRPARYAVA-REAVGVAPLTAHGTLLVNDVLAACVAVLESHQW 352
DB 299 QRYVVAERGGDRRLPAVHVSVTLREEGAGVAPLTAHGTLLINRVLAACVAVLESHQW 358
QY 353 AHRAFAPLRLIHA-LGALLP-----GNAV-----QPT-GMHVYSRLLYELA 391
DB 359 AHRAFAPLRLIHALAALAPARTDGGGGSIPAAQATARGAEPTAGIHWYSQLLYHIG 418
QY 392 EELL 395
DB 419 TWLL 422

RESULT 6
PCT-US95-15923-20
Sequence 20, Application PC/TUS9515923
GENERAL INFORMATION:
APPLICANT: The Johns Hopkins University School of Medicine, et al.
TITLE OF INVENTION: NOVEL HEDEHOG-DERIVED POLYPEPTIDES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: U.S.A.
ZIP: 92037
MEDIUM TYPE: floppy disk
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/15923
FILING DATE: 04-DEC-1995
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/043W01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-15923-20

Query Match 58.1%; Score 1214; DB 4; Length 437;
Best Local Similarity 58.5%; Pred. No. 3.5e-126;
Matches 248; Conservative 54; Mismatches 82; Indels 40; Gaps 12;

QY 7 LPLCLALLA-----LPAQSCGPGGPGVGRRRYARKOLVPLLYKQFVGYVPTIGASG 61
DB 4 LARCELYLASSLLVCPGLACGPRG-FGKRHPKK-LTPLAYKQFIPVAEKTIGASG 61
QY 62 PAGRVARGSERFDLVPNYNPDIIFKDEENSGADRLMTERCKERYNALAIAMNMPGV 121
DB 62 RYGRKTRNSERKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAIAMNMPGV 121
QY 122 RLRTGEMWEDGHHAQDSLYEGRALDITTSRDNRKYGILLARLAVEGFDWVYESRNH 181
DB 122 KLRTVEGWEDGHSESLHYEGRAVDITTSRDNRKYGILLARLAVEGFDWVYESRNH 181
QY 182 VHVSVKADNSLAVRAGCGPGNATVRLMSGERKGLRELHNGDMVLAADASGRVPTPYLL 241
DB 182 IHCSVAENSVAAGSGCGPGSATVHLEGGTKLVKDLRPGDVLADDOGRLLYSDFLT 241
QY 242 FLDRDLQRRASFVAVETEMPRLKLLTPMHLVFAA---RGPAAPGDFAPVAFARLRAG 297
DB 242 FLDRDGAQKRVYVETLEPRERLLTAHLFLVAPHNDGPTPGP---SALFASRVRG 298
QY 298 DSVLA---PGGD-ALRPARYAVA-REAVGVAPLTAHGTLLVNDVLAACVAVLESHQW 352
DB 299 QRYVVAERGGDRRLPAVHVSVTLREEGAGVAPLTAHGTLLINRVLAACVAVLESHQW 358
QY 353 AHRAFAPLRLIHA-LGALLP-----GNAV-----QPT-GMHVYSRLLYELA 391
DB 359 AHRAFAPLRLIHALAALAPARTDGGGGSIPAAQATARGAEPTAGIHWYSQLLYHIG 418
QY 392 EELL 395
DB 419 TWLL 422

RESULT 7
PCT-US95-02315-2
Sequence 2, Application PC/TUS9502315
GENERAL INFORMATION:
APPLICANT: Jessell, Thomas M.
APPLICANT: Dodd, Jane
APPLICANT: Roelink, Henk
APPLICANT: Edlund, Thomas
TITLE OF INVENTION: DNA ENCODING A VERTEBRATE HOMOLOG OF
TITLE OF INVENTION: HEDEHOG, VHH-1, EXPRESSED BY THE NOTOCHORD, AND USES
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/02315
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: John P. White
 REGISTRATION NUMBER: 28,678
 REFERENCE/DOCKET NUMBER: 45375-A-PCT
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 278-0400
 TELEFAX: (212) 391-0525
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 437 amino acids
 TYPE: amino acid
 TOPOLOGY: 1linear
 MOLECULE TYPE: protein
 PCT-US95-02315-2

Query Match	57.68;	Score 1203;	DB 4;	Length 437;
Best Local Similarity	57.88;	Pred. No. 5.7e-125;		
Matches 245;	Conservative 52;	Mismatches 87;	Indels 40;	Gaps 10

```

QY 7 LLLPCLC-----ALLALPAOCSGGRBPVGRRRRAKQVLYLKQVPGPPTLGCASG 61
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Dd 4 LLANGELVALLASSLLVCGCLGACGPRG-FGKRQHPK-LTPLAQKQIPVNAETTLGASG 61
QY 62 PAEGRVAGSERFPDLPVNPNDIIFKDEENSGADRLTERCKERYNALIAVNMMPGV 121
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Dd 62 RYBEKRTIENSSEFPELTPNYPNDIIFKDEENTGADRLMTQCKDKLNALIASVNMQMPGV 121
QY 122 RLRTYEGMDEGCHHAOSLIHESGRALDITTSRDNRNKYGLIARLAVAGSGDWYEGSRNH 181
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Dd 122 KLRTYEGMDEGCHHSEELTEGRAVDITTSRDNRSKYGLIARLAVAGSGDWYTESKAR 181
QY 182 VHSVVKADNSLAVRAGGCFPGNATVRLTMSGERKGLREIHRGDWVLADAGSRVVPYVL 241
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Dd 182 IHCSVKAKENSVAAKSDGCFPGSATVHLEGGCTKLYKDLSPEDRYLADDDGRLLYSDFLT 241
QY 242 FLDDDLQRRASFVAVETWPRKLLTPMHLVFA---KGPAPDGPAPVFAARRLRAG 291
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Dd 242 FLDDDEGAKKVFYVIEPREPRERLLTAHLLFPAHPNDSGPPRGP---SELFASRRVPG 291
QY 298 DSVLA---PGGD-ALRPARVARVA-REEVAVGFAPLTAHGTLVNDVLASCYAVLSESHW 351
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Dd 298 QRVYVVAERGGDRLLPAAVHVSYTLREEBACATYAPLIRADGITLNNRYLASCAVIEEHSW 351
QY 353 AHRFAFAPRLRLHA-LGALLPG-----GAVQPTGMHWYSLRLYLRLA 391
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Dd 359 AHRFAFAPRLHAALLALAPARTDGGGGGSIAPQSVAEARGAGPAPGIMHYSQLYHIG 418
QY 392 EELL 395
| |
Dd 419 TWLL 422

RESULT 8
US-08-176-4278-2
; Sequence 2, Application US/08176427B
; Patent No. 5789543
; GENERAL INFORMATION:
; APPLICANT: Ingham, Phillip W.
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Tabin, Clifford J.
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
; TITLE OF INVENTION: Proteins and Uses Related Thereo
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:

```

```

1 ADDRESS: LAHIVE & COCKFIELD
2 STREET: 60 State Street
3 CITY: Boston
4 STATE: MA
5 COUNTRY: USA
6 ZIP: 02109
7 COMPUTER READABLE FORM:
8 MEDIUM TYPE: Floppy disk
9 COMPUTER: IBM PC compatible
10 OPERATING SYSTEM: PC-DOS/MS-DOS
11 SOFTWARE: ASCII(text)
12 CURRENT APPLICATION NUMBER:
13 APPLICATION NUMBER: US/08/176,427B
14 FILING DATE: 30-DEC-1993
15 CLASSIFICATION: 435
16 ATTORNEY/AGENT INFORMATION:
17 NAME: Vincent, Matthew P.
18 REGISTRATION NUMBER: 36,709
19 REFERENCE/DOCKET NUMBER: HWI-006
20 TELECOMMUNICATION INFORMATION:
21 TELEPHONE: (617) 227-7400
22 TELEFAX: (617) 227-5941
23 INFORMATION FOR SEQ ID NO: 2:
24 SEQUENCE CHARACTERISTICS:
25 LENGTH: 425 amino acids
26 TYPE: amino acid
27 TOPOLOGY: linear
28 MOLECULE TYPE: protein
29 US-08-176-427B-2

```

Query Match	57.4%	Score 1198	DB 1	Length 425
Best Local Similarity	59.4%	Pred. No. 2e-124		
Matches 244	Conservative 56	Mismatches 91	Indels 20	Gaps 10

```

OY      1 MALLTNLP---LCCMLALPAQSCGPGRGPRRRYARKOLVPLLYKQFVPCVPERTL 57
      1 | | | | : | | | : | | | | : | | | : | | | | : | | | : | | | |
Db      4 MLLTLRILVGFIC--ALVYSSGLTCGPGRG--IGKRHPKK--LTLPLAYKQFIRVNAEKT 59
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      58 GASGPAEGRVARGSERFERDLPVNPNDPIFKDEENSGADRLMTERCKERVNALAIAYMM 117
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      60 GASGYEKKITRNSERPEFKELTPNPNDPIFKDEENTGADRLMTQCKDKLINALAISYMQ 119
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      118 WPGYRLRNTBECMDDDGHHADQSLYBEGRALDITTSDDNRKKYGLLALAYAGEDWYYE 177
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      120 WPGYRLRNTBECMDDDGHHSEESLYBERAVDITTSDDNRKSYGLRLAYEAGDWYYE 179
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      178 SRNHVSVKADNSLAVRAGCEFGNATVRLMSGERKLELHNGDWVLAADASGRVPT 237
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      180 SKAHTHCYKAEVNAKSSGCGCFSGSATVHLHEGCTLVYNDLSBGDRVLAADADGRLLYS 239
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      238 PVLLEFDRLDQRRASFVAVETEMPRLKLLTPMHLVFA--RCGPAPAGDFA--PVFARR 293
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      240 DELFELDMDSRKLFEVIEROPRRLTLTAALLFLVAPDHQNSFATSGTSGCALPASN 299
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      294 LRAGDS--VLAPGDALRPANAVVA--REAVGYAFELTAHGTLLVNDVLASCAYVESH 350
      : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      300 VKPGRYVYLEEGQQLLPASVSHVSLEKASGAYAPLTAGOTLLIRVLAASCAYVEEH 350
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      351 QWARRAFAPLRLIHA--LGALLPGGAV-----QPTGMHYSRLYRLAEELL 395
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      360 SWAHNAFAPFLRAGLLAALCPDGAIPTAATTTGIMYSRLYRISSWVL 410
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT      9
US-08-356-060A-8
; Sequence 8, Application US/08356060A
; Patent No. 5844079
;
GENERAL INFORMATION:
; APPLICANT: Ingham, Phillip W.
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Tabin, Clifford J.
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
;

```


Db 418 PGADAPGAGATAGIHWSQLTYOIGTWLL 447

RESULT 11
US-08-748-591-9

; Sequence 9, Application US/08748591

; Patent No. 5759811

; GENERAL INFORMATION:

; APPLICANT: Epstein, Ervin

; APPLICANT: Hu, Zhilan

; APPLICANT: Bonifas, Jeanette

; TITLE OF INVENTION: Mutant Human Hedgehog Gene

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish and Richardson

; STREET: 2200 Sand Hill Road

; CITY: Menlo Park

; STATE: CA

; COUNTRY: USA

; ZIP: 94025

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/748,591

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Sherwood, Pamela J

; REGISTRATION NUMBER: 36,677

; REFERENCE/DOCKET NUMBER: 06510/067001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 322-5070

; TELEFAX: (415) 854-0875

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 462 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-748-591-9

Query Match 56.4%; Score 1177; DB 1; Length 462;

Best Local Similarity 54.4%; Pred. No. 4.8e-122;

Matches 245; Conservative 54; Mismatches 93; Indels 58; Gaps 10;

Db 1 MALLTNLPCLCL-ALLALPAQSCGPGRGVGRRRYARKOLVPLLYKQFPGVPERLGA 59
1 MLLARCLLVVSSLVCSGLACGPRG-FGKRHRPK-LTPLAYKQFIPNAEKTIGA 58
QY 60 SGAPEGVANGSERFRLVNPNDIIFKDENSGADRLMTERCKERNALAIAMNMP 119
119 GVKLRVTEGDEGHSHSEELHYEGRAVDITTSRDRSKYGMALARLAVEAGFDWVYYESK 178
Db 119 GVKLRVTEGDEGHSHSEELHYEGRAVDITTSRDRSKYGMALARLAVEAGFDWVYYESK 178
QY 180 NHVAVSKADNSLAVRAGGCPGNATVRLMSGERKGLRELRGDMVLAADAGSVPTPV 239
179 AHICSTKAENSVAAGSGGCPGSAIVHLDEGGTKLVKDLSPGRVLAADQGRLLSDF 238
Db 240 LFLDRLQRRASFVAETEMPRKLLLTTPMHLVFAA-----RGPAPAG 284
239 LTFDRLDQAKKVFYVETREPRERLLTAAHLFVAPHNDSATGEPRASSGSP-PSGG 297
QY 285 DEAP--VFARLRAGDSVLA-----PGDALPRAVARA-REEVGVFAPLTAHGTLLVN 337

Db 298 ALCPRLAFASRVPRGQRYVVAERDGRLLPAAVHSVTISEAAQAYADLIAQGTILLN 357

QY 338 DYLAACVAVLESHQMAHRAFAPLRLHALGALL----- 371

Db 358 RVLASCAVIEEHSMAHRAFAPLRLHALGALLAALAPARTDNGSGGGGDRGGGGRVALTA 417

QY 371 PGAVOP-----TGMHYSRLYRLAEELL 395

Db 418 PGADAPGAGATAGIHWSQLTYOIGTWLL 447

RESULT 12

US-08-356-060A-13

; Sequence 13, Application US/08356060A

; Patent No. 5844079

; GENERAL INFORMATION:

; APPLICANT: Ingham, Phillip W.

; APPLICANT: McMahon, Andrew P.

; APPLICANT: Tabin, Clifford J.

; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing

; NUMBER OF SEQUENCES: 47

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD

; STREET: 60 State Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII(text)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/356,060A

; FILING DATE: 14-DEC-1994

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/176,427

; FILING DATE: 30-DEC-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Vincent, Matthew P.

; REGISTRATION NUMBER: 36,709

; REFERENCE/DOCKET NUMBER: HMI-006CP

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 227-7400

; TELEFAX: (617) 227-5941

; INFORMATION FOR SEQ ID NO: 13:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 475 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-356-060A-13

Query Match 56.4%; Score 1177; DB 2; Length 475;

Best Local Similarity 54.4%; Pred. No. 5e-122;

Matches 245; Conservative 54; Mismatches 93; Indels 58; Gaps 10;

Db 1 MALLTNLPCLCL-ALLALPAQSCGPGRGVGRRRYARKOLVPLLYKQFPGVPERLGA 59
1 MLLARCLLVVSSLVCSGLACGPRG-FGKRHRPK-LTPLAYKQFIPNAEKTIGA 58
QY 60 SGAPEGVANGSERFRLVNPNDIIFKDENSGADRLMTERCKERNALAIAMNMP 119
119 GVKLRVTEGDEGHSHSEELHYEGRAVDITTSRDRSKYGMALARLAVEAGFDWVYYESK 178
Db 119 GVKLRVTEGDEGHSHSEELHYEGRAVDITTSRDRSKYGMALARLAVEAGFDWVYYESK 178
QY 120 GVKLRVTEGDEGHSHSEELHYEGRAVDITTSRDRSKYGMALARLAVEAGFDWVYYESK 179
119 GVKLRVTEGDEGHSHSEELHYEGRAVDITTSRDRSKYGMALARLAVEAGFDWVYYESK 178

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2000, 07:29:48 ; Search time 1833.88 Seconds

(without alignments)
-631.772 Million cell updates/sec

Title: US-08-900-220-8

Perfect score: 1191
Sequence: 1 ATGCTCTCTGACCAATCT.....CGAGAGACTACTGGCTGA 1191

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl:
1: gb_dal:*
2: gb_dal:*
3: gb_om:*
4: gb_om:*
5: gb_pat:*
6: gb_pat:*
7: gb_pat:*
8: gb_pat:*
9: gb_pat:*
10: gb_pat:*
11: gb_pat:*
12: gb_pat:*
13: gb_pat:*
14: gb_pat:*
15: gb_pat:*
16: gb_pat:*
17: gb_pat:*
18: gb_pat:*
19: gb_pat:*
20: gb_pat:*
21: gb_pat:*
22: gb_pat:*
23: gb_pat:*
24: gb_pat:*
25: gb_pat:*
26: gb_pat:*
27: gb_pat:*
28: gb_pat:*
29: gb_pat:*
30: gb_pat:*
31: gb_pat:*
32: gb_pat:*
33: gb_pat:*
34: gb_pat:*
35: gb_pat:*
36: gb_pat:*
37: gb_pat:*
38: gb_pat:*
39: gb_pat:*
40: gb_pat:*
41: gb_pat:*
42: gb_pat:*
43: gb_pat:*
44: gb_pat:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	968.6	81.3	1191	12 MMDH	X76292 M.musculus
2	967.6	81.2	1190	5 AR021199	AR021199 Sequence
3	967.6	81.2	1190	5 AR063081	AR063081 Sequence
4	630.4	55.0	344025	51 AR011603	AR011603 Homo sapi
5	620.0	52.6	6200	9 AR010994	AR010994 Homo sapi
6	462	38.8	1958	4 GGU58511	U58511 Gallus galli
7	429.4	36.1	2103	12 MMU85610	U85610 Mus musculus
8	419.8	35.2	1313	5 AR021201	AR021201 Sequence
9	419.8	35.2	1313	5 AR063083	AR063083 Sequence
10	418.2	35.1	1314	12 MMSRH	X76290 M.musculus
11	415.6	34.9	1197	4 XU026350	U26350 Xenopus lae
12	407	34.2	1715	12 RATVH1X	L27340 Rat (Vh-1)
13	406.6	34.1	1191	4 XU026349	U26349 Xenopus lae
14	390.6	32.8	1425	5 AR063085	AR063085 Sequence
15	390.6	32.8	1576	9 HIMS5H	L38518 Homo sapien
16	385.2	27.3	1635	4 AR047466	AR047466 Notophila
17	381.4	32.0	1277	5 AR021198	AR021198 Sequence
18	381.4	32.0	1277	5 AR063080	AR063080 Sequence
19	381.4	32.0	1567	4 CHK2PAMED	L28099 Gallus galli
20	373	31.3	1718	4 AB029748	AB029748 Parolicht
21	370.2	31.1	1807	4 NEMSPA	D63339 Cynops pyrr
22	357.4	30.0	1011	12 MMTIR	X76291 M.musculus
23	355.8	29.9	1056	5 AR021200	AR021200 Sequence
24	355.8	29.9	1056	5 AR063082	AR063082 Sequence
25	350.6	29.4	1549	4 AF003532	AF003532 Pleurodel
26	335.6	28.2	939	5 AR063086	AR063086 Sequence
27	335.6	28.2	1277	9 HUMTIR	L38517 Homo sapien
28	335.6	28.2	1277	13 G28584	G28584 human STS S
29	316.4	26.6	1251	4 DR030710	U30710 Danio rerio
30	316.2	26.5	1256	5 AR021202	AR021202 Sequence
31	316.2	26.5	1256	5 AR063084	AR063084 Sequence
32	316.2	26.5	1684	4 BRSONHP	L23569 B.rerio shh
33	316.2	26.5	2600	4 ZEPVH1A	L27585 Brechydanto
34	314.6	26.4	1257	4 DR030711	U30711 Danio rerio
35	308.8	25.9	2447	4 XELXSH	L39213 Xenopus lae
36	308.4	25.9	386	12 AF148226	AF148226 Rattus no
37	307.2	25.8	1335	4 XU026314	U26314 Xenopus lae
38	305.2	25.6	1230	4 XU026404	U26404 Xenopus lae
39	304	25.5	1479	4 DRECHTDA	Y08426 D.rerio MRN
40	301.8	25.3	1491	9 AB018401	AB018401 Homo sapi
41	301.4	25.3	303	9 AB010581	AB010581 Homo sapi
42	301.4	25.3	2305	34 S66384	S66384 hh-segment
43	301.4	25.3	3091	34 DROHHA	L02793 Drosophila
44	299.8	25.2	2046	34 DROHNB	L05404 Drosophila
45	297.8	25.0	1546	4 XELVH1A	L35248 Xenopus lae

ALIGNMENTS

[illegible]

QY	181	GGCGACGGAGGGAGGGAGGGGCGCAAGGGGCTCCGACGGCTTCGGGAGACTGTGGCCAAC	240
Db	181	GGCGACGGAGGGAGGGAGGGGCTCGACAGGGGCTTCGGGAGACTGTGACCCAAC	240
QY	241	TACAAACCCCGACATCTTCAAGGATGAGAGAACAGTGGAGACCGACCGCTGATGACC	300
Db	241	TACAAACCCCGACATATCTTCAAGGATGAGAGAACACGGGGGCGACACCGCTGATGACA	300
QY	301	GAGGTTGCAAGAGAGAGGTGAACGCTTTTGCCATTGCCGTATGAACATGTGGCCGGA	360
Db	301	GAGGTTGCAAGAGAGCGGGTGAACCGCTCTAGCCATGCCGTATGAACATGTGGCCGGA	360
QY	361	GTGGCCCTACGAGTGTGAGGGCTGGAGACGAGAGAGGCGCACAGCGTCAAGATTCACTC	420
Db	361	GTAGCCCTACGTTGACTGAAGGCTGGAGAGAGAGCGCACACGACACAGGATTCACTC	420
QY	421	CACATACGAAGGCCCTGCTTTGAGCATCACTACGTCTGACCGCGACCGGACACAAATGATGG	480
Db	421	CACATACGAAGGCCGTGCTTTGAGCATCACTACAGCTCTGACCGGACCGGATTAATGATGT	480
QY	481	TTGCTGGCGGCGCTCGCAGTGAAGCCGCGCTTGACTGGGTCTACTACGATCCCGCAAC	540
Db	481	TTGTGGGCGCCCTCACTGCTGGAAGCCGCGATTGATGATGGGTCTACTACGAGTCCCGCAAC	540
QY	541	CACGTCCACAGTGTGGTCAAGAGCTGATTAATCACTAGTGGCGGTTCGGGGCGCGCGCTT	600
Db	541	CACATCCACAGTATGGTCAAAAGCTGATTAATCACTAGTGGCGGTTCGGAGCGGAGCGCTTT	600
QY	601	CCGGGAATGCAATCTGCCCCCTGTGGAGCGCGGACCGGAAGAGGCTCGGGAACTGCAC	660
Db	601	CCGGGAATGCAAGGTGGCGCTTGGAGACCGCGGACCGGAAGAGGCTCTAGGGAATCAAT	660
QY	661	CGCGAGACTGGGTTTTTGGCGCGGATGCGTCAAGCGCGGGTGGTCCACGCGCGTCTG	720
Db	661	CGTGTGACTGGTACTGGAGTGGCGCGCTGATGACAGCGGGCGGAGTGGTACCACGCGCAATGCTG	720
QY	721	CTCTCTCTGGAACGGGACTTGGACAGCGCGCGGCTTATTTGTGTGGCTGTGGAACGCAATGG	780
Db	721	CTCTCTCTGGAACCGGGATCTGACAGCGCGCGCTCTTGTGTGGCTGTGGAACCGCAATGG	780
QY	781	CCCTCAGCAGCAACGTGTGCTCACGCGCTGGACACTGGTGTGGCGCTGAGAGGGCGGCG	840
Db	781	CCTCCGCGCAAACTGTGTGCTCACACCTTGGCATCTGTGTGTGCTGTGCGAGGCGCAAGG	840
QY	841	CCCGCGCAGGCGCACTTGGCACCGGATGTTCCGCGCGCGCTCAAGCGCTGAGGAGACTCGGTG	900
Db	841	CCGTGCTCAGGTGACTTGTGACCGGGTGTTCGCGCGCGCTTACGTACGTGCGGACTCGGTG	900
QY	901	CTGGCGCCCGCGGGGATGCGCTTGGGCCACAGCGCGCGTGTGCCCCGTGTGGCGGGAGAA	960
Db	901	CTGGCTCCCGCGGGGAGCGGCTGCACAGCGCGCGGTGTGCCCCGTGTGGCGGGAGAA	960
QY	961	GCGCGGGCGGTTCGCGCGCGGCTCACCGGGGCAAGGGAGCGCTGTGTAAAGATGCTCTG	1020
Db	961	GCGGTGGCGGTTCGCACTCCCTCACCTGGACAGGGAGCGCTGTGTAAAGATGCTCTC	1020
QY	1021	GCTCTTGTACTACGGGTTGTGAGAGTCAACAGTGGGCGCACCGCGCTTTTGCCCGCTTG	1080
Db	1021	GCTCTCTGTACAGGGGTTCTTAGAGATCAACAGTGGGCGCACCGCGCTTGCCCGCTTG	1080
QY	1081	AGACTGCTGCAGCGGCTAGGGGGGCTGCTCCCGCGGGGGCGGTCCACAGCACTGGCATG	1140
Db	1081	CGCGCTGCTGCAGCGGCTAGGGGGGCTGCTCCCGGGGGGTGCAAGTCCACCGCACTGGCATG	1140
QY	1141	CATTGTGACTCTCGGCTCTCTACCGCTTAAGCGGAGAGAGCTACTGGGCTGA	1191
Db	1141	CATTGTGACTCTCGCTCTCTTACCGCTTGGCGGAGAGAGTAAATGGGCTGA	1191
RESULT 2			
AR021199 1190 bp DNA PAT 05-DEC-1998			
LOCUS AR021199 1190 bp DNA PAT 05-DEC-1998			
DEFINITION Sequence 3 from patent US 5789543.			

[illegible]

QY	781	CCTCAGCAAACTGTTGCTACAGCCCGGAGACCTGGGTTTGGCCGCGAGAGGCGGAGC	840
QY	781	CCTCAGCAAACTGTTGCTACAGCCCGGAGACCTGGGTTTGGCCGCGAGAGGCGGAGC <td>840</td>	840
Db	781	CCTCAGCAAACTGTTGCTACAGCCCGGAGACCTGGGTTTGGCCGCGAGAGGCGGAGC	840
QY	841	CCCGGCGCAAGGGAATTTGTCACACCGGTGTTCGGGCGCCGCGCTACGCGCTGGGAGACTCGGTG	900
Db	841	CCCGGCGCAAGGGAATTTGTCACACCGGTGTTCGGGCGCCGCGCTACGCGCTGGGAGACTCGGTG	900
QY	901	CTGGGCGCCCGCGGGGAGATGCGCTTGGGCCAGCGCGCGCTGGGCCGCGGGAGAA	960
Db	901	CTGGGCGCCCGCGGGGAGAGCGCGCTCAGCGCGCGCGCTGAGCCCGGTGGCGCGAGAA	960
QY	961	GCGGCGGGGGGTTTGGCGCGCGCTACCGCGGACCGGAGAGCTGCTGTAGACGATTCCTG	1020
Db	961	GCGGCGGGGGGTTTGGCGCGCGCTACCGCGGACCGGAGAGCTGCTGTAGACGATTCCTG	1020
QY	1021	GCGCTTGTGTACGCGGTTCTGTGAGAGTACACAGTGGGCGCACCGGCTTTTGGCCCTTG	1080
Db	1021	GCGCTTGTGTACGCGGTTCTGTGAGAGTACACAGTGGGCGCACCGGCTTGTGGCCCTTG	1080
QY	1081	AGACTGCTGACAGCGGCTTAGGGGCGCTGCTCCCGCGCGGCGCGCTTCCAGCTGAGCTG	1140
Db	1081	CGGCTGCTGACAGCGGCTTAGGGGCGCTGCTCCCGCGGCGCGCTTCCAGCTGAGCTG	1140
QY	1141	CATTGTACTCTCGGCTCCTTACCGCTTACCGCTTACCGCTTACCGCTTACCGCTT	1190
Db	1141	CATTGTACTCTCGGCTCCTTACCGCTTACCGCTTACCGCTTACCGCTTACCGCTT	1190
RESULT	3		
LOCUS	AR063081	1190 bp	DNA
DEFINITION	Sequence 2 from patent US 5844079.		PAT
ACCESSION	AR063081		
VERSION	AR063081.1	GI:5990772	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 1190)		
AUTHORS	Ingham, P.W., McMahon, A.P. and Tabin, C.J.		
TITLE	Vertebrate embryonic pattern-inducing proteins, and uses related thereto		
JOURNAL	Patent: US 5844079-A 2 01-DEC-1998;		
FEATURES	Location/Qualifiers		
source	1..1190		
BASE COUNT	194 a 371 c 399 g 226 t		
ORIGIN			
Query Match	81.2%;	Score 967.6;	DB 5;
Best Local Similarity	88.3%;	Fred. No. 7.9e-153;	
Matches 1051;	Conservative	0;	Mismatches 139; Indels 0; Gaps 0;
QY	1	ATGGCTCTCTGTACCAATCTACTGCGCTTGTGCTTGGGCACTTGTGGCGCTGCGAGCC	60
Db	1	ATGGCTCTCTGTACCAATCTACTGCGCTTGTGCTTGGGCACTTGTGGCGCTGCGAGT	60
QY	61	CAGAGCTGGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG	120
Db	61	CAGAGCTGGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG	120
QY	121	GTCGCGCTCTCTTACAGCAATTTGTGCTGCGGCGTGGCAGAGCGGACCTTGGGCGCGAGT	180
Db	121	GTCGCGCTCTCTTACAGCAATTTGTGCTGCGGCGTGGCAGAGCGGACCTTGGGCGCGAGT	180
QY	181	GGGCGCAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	240
Db	181	GGGCGCAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	240
QY	241	TACACCCCGACATATCTTCAAGGATGAGGAAACAGTGAAGCGGACCGCTGATGAC	300

Db	241	TACAA	CCCCCGCATATATCTTCAAGAGATGAGAGAAACAGCGCGCGCAGACCGCCTGATGACA	300
Qy	301	GAGCGTTGCAAGAGAGAGGGTGAACGCTTTTGGCCATTGCGCTGATGAACATGTGGCCCGGA	360	
Db	301	GAGGGTTGCAAAAGAGCGGGGTGAACGGCTCTAAGCCATGGCGGTGATGAACATGTGGCCCGGA	360	
Qy	361	GTCGCGCTACAGAGTATGAGGGCTGGAGACGAGAGCGGCACACAGCTATGAGATTCACTC	420	
Db	361	GTACGCTACTGCTGATCTGAAGAGCTGGAGACGAGAGAGCGGCACACGACGACAGATTCACATC	420	
Qy	421	CACATACGAAGCCGCTGCTTTTGGACATCACTACTAGCTTGACCCGACCGGCAACAATATGGG	480	
Db	421	CACATACGAAGCCGCTGCTTGGCTTGGAGAAAGCCGATTGATGATGGGTCTACTACGATGTCGGCAAC	480	
Qy	481	TTGTGGCGCGCCCTCGGACAGTGAAGGAAGCCGGTTCGATGGGTCTACTAGATGATCGCGCAAC	540	
Db	481	TTGTGGCGCGCCCTAGTGTGGAGAAAGCCGATTGATGATGGGTCTACTACGATGTCGGCAAC	540	
Qy	541	CACGCTCCAGCTGTGCGTCAAAAGCTGATATATCACTAGCGCGCTCCGGCGCGCGCTGCTTT	600	
Db	541	CACATCCACGATATGATGCTCAAAAGCTAATCACTAGTGGCGGTCCGAGCGGAGGCTGCTTT	600	
Qy	601	CCGGGAATATGCACTGTGCGCTGTGTGAGCGCGGACCGGAAGGGCTGCGGGACTGCAC	660	
Db	601	CCGGGAATATGCACTGATGCGCTCTTCGGAAGCGGGAAGGAAGGGCTGTGAGGAATCTACAT	660	
Qy	661	CGCGAGACATGGGTTTGGGGCGCGAATGCGTCAAGCGCGGGTGGTGGCCAGCGCGGTGCTG	720	
Db	661	CGTGTGATCTGGTACTGAGCGCGCTGATGACAGCGGGCGGAGTGGTATCCACGCGCACTGCTG	720	
Qy	721	CTCTTCCTGGAACCGGACCTTGACAGCGCGCGGGCTCTCATTTGTGGCTGTGAGACAGCAATGG	780	
Db	721	CTCTTCCTGGAACCGGAGATGACAGCGCGCGGGCTCTCTGTGGCTGTGAGACAGCAAGCGG	780	
Qy	781	CCTCCAGCAGCAACCTGTGCTACAGCCCTGGCACCTGTGTGGCGCTGAGAGCGCGCGG	840	
Db	781	CCTCCGCGCAAACTGTGCTACACACCTTGGCACTGTGTGGCTCTGCGCGGCGACACG	840	
Qy	841	CCGCGCGCAGCGCACTTGGACCGGATTTGGCGCGCGCGGCTACGCGCTGGGAGACTCGGATG	900	
Db	841	CCTGCTCCAGGTATGCTTGGCACCGCGATTTGGCGCGCGCTTACGTGCTGTGGCACTCGGATG	900	
Qy	901	CTGGCGCCCGCGGGGATTCGCTTGGCGCACGCGCGCTGGGCCCTGTGTGGCGCGGAGANA	960	
Db	901	CTGGCTCCCGCGGGGAGACGCGCTTCACACCGCGCGCGTGAACCGCGCTGGCGCGGAGANA	960	
Qy	961	GCGCGGGCGGCTTCCGCGCGGCTCACCGCGGACGAGGAACGTGTGGTGAACAATGCTCTG	1020	
Db	961	GCGCGGGCGGCTTCCGACCGGCTTACGCGCGGACGAGGACGCTGTGGTCAACAACGTCTCTC	1020	
Qy	1021	GCGCTTGTCTACGCGGCTTGTGAGAGTACACACATGGGCGCACGCGCTTTTCCGCCCTTGG	1080	
Db	1021	GCGCTGTCTACGCGGCTTGTGAAGAGTACACACATGGGCGCACGCGCTTTTGGGCCCTTGG	1080	
Qy	1081	AGACTGCTGACAGCGGCTTAGAGGGGCGCTGCTCCCGCGGGGCGGTCCAGCTGCATGCGATG	1140	
Db	1081	CGGCTGCTGACAGCGGCTTCCGCGGCTGCTCTCCCTGGGGGTGACAGTCCAGCCGACTGGCATG	1140	
Qy	1141	CATTGTACTCTGGGCTGCTTACACCGGTTAGGGGAAGAGCTACTGGCGTG	1190	
Db	1141	CATTGTACTCTGGCTCTCTTACTACCGGTTAGGGGAAGAGTAAATGGCGTG	1190	

ORGANISMISM

REFERENCE AUTHORS

Homo sapiens
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Mammalia:
Eutheria: Primates: Catarrhini: Hominoidea: Homo.
1 (bases 1 to 241275)
Muzny, D. M., Adams, C., Bailey, M., Barbara, J., Blankenburg, K.,

TITLE
JOURNAL
REFERENCES
AUTHORS
TITLE
JOURNAL

Submitted (08-OCT-1999) Human Genome Sequencing Center, Department
 Direct Submission
 Worley, K. G.
 2 (bases 1 to 241275)
 Unpublished
 Direct Submission

COMMENT

On Feb 19, 2000 this sequence replaced g1:6728920.

Center: Baylor College of Medicine

Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Inform

Center clone name: RP11-386G11

Sequencing vector: M13; L08821

Chemistry: Dye-terminator bld Dye: 99% of reads

Assembly program: Phitap; version 0.980611
 Compiling qual4tw; 166379 bytes at 10ast

Consensus quality: 189167 bases at least
Consensus quality: 300038 bases at least

Estimated insert size: 207648; sum-of-cor

Qualitv coverage: 3.4x in Q20 bases; agardose-1p estimation

* NOTE: This is a 'working draft' sequence. It currently

* is not known and their order in this sequence record is

* runs of N but the exact sizes of the gaps are unknown

* This record will be updated with the finished sequence as soon as it is available and the accession number will

775: cont'd of 775 ha in length

```

*      /93: gap of unknown length
*      796
*      1619: cont'd of 824 bn in length

```

1020	1039: gap of unknown length
1640	2534: cont'd of 895 ha in length

2333	2334:	gap of unknown length
2555	3411:	constit of 857 bp in length

```

*      3412      3431: gap or unknown length
*      3432      4198: cont'd of 767 bp in length

```

```

*      4199      4218: gap of unknown length
*      4218      5149: cont'd of 921 bc in length

```

3130 3109: gap of unknown length

FEATURES	SOURCE
exon	1. .626 /gene="hedgehog" /number=3 /start=200 /end=236 /strand=+
gene	join(AB010581.1:1. .303, AB010993.1:1. .262, 1. .626) /gene="hedgehog" /start=200 /end=236
CDS	1. .626 /gene="hedgehog" /start=200 /end=236
REFERENCE	Submitted (03-FEB-1998) to the DDBJ/EMBL/GenBank databases. Genshu Tate, Shoin University Fujigakka Hospital, Department of Surgical Pathology, Fujigakka 1-30, Aoba-ku, Yokohama, Kanagawa 227-8501, Japan (E-mail:six10352@super.win.or.jp, Tel:81-45-974-6632, Fax:81-45-972-6242)
ORGANISM	Location/Qualifiers 1. .626 /organism="Homo sapiens" /db_xref="taxon:9606" join(AB010581.1:1. .303, AB010993.1:1. .262, 1. .626) /gene="hedgehog" /start=200 /end=236
LOCUS	AB010581s3 626 bp DNA PRI 07-JAN-2000
DEFINITION	Homo sapiens hedgehog gene, exon 3 and complete cds.
ACCESSION	AB010994
VERSION	AB010994.1 GI:287455
KEYWORDS	hedgehog.
SEGMENT	3 of 3
SOURCE	Homo sapiens DNA. Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	Tate, G., Kishimoto, K. and Mitsuura, T. Expression of Sonic hedgehog and its receptor Patched/Smoothed in human cancer cell lines and embryonic organs J. Biochem. Mol. Biol. Biophys. 4, 27-34 (2000)
AUTHORS	Tate, G., Endo, Y. and Mitsuura, T.
REFERENCE	Submitted (03-FEB-1998) to the DDBJ/EMBL/GenBank databases. Genshu Tate, Shoin University Fujigakka Hospital, Department of Surgical Pathology, Fujigakka 1-30, Aoba-ku, Yokohama, Kanagawa 227-8501, Japan (E-mail:six10352@super.win.or.jp, Tel:81-45-974-6632, Fax:81-45-972-6242)
AUTHORS	Submitted (03-FEB-1998) to the DDBJ/EMBL/GenBank databases. Genshu Tate, Shoin University Fujigakka Hospital, Department of Surgical Pathology, Fujigakka 1-30, Aoba-ku, Yokohama, Kanagawa 227-8501, Japan (E-mail:six10352@super.win.or.jp, Tel:81-45-974-6632, Fax:81-45-972-6242)
REFERENCE	Submitted (03-FEB-1998) to the DDBJ/EMBL/GenBank databases. Genshu Tate, Shoin University Fujigakka Hospital, Department of Surgical Pathology, Fujigakka 1-30, Aoba-ku, Yokohama, Kanagawa 227-8501, Japan (E-mail:six10352@super.win.or.jp, Tel:81-45-974-6632, Fax:81-45-972-6242)
AUTHORS	Submitted (03-FEB-1998) to the DDBJ/EMBL/GenBank databases. Genshu Tate, Shoin University Fujigakka Hospital, Department of Surgical Pathology, Fujigakka 1-30, Aoba-ku, Yokohama, Kanagawa 227-8501, Japan (E-mail:six10352@super.win.or.jp, Tel:81-45-974-6632, Fax:81-45-972-6242)
REFERENCE	Submitted (03-FEB-1998) to the DDBJ/EMBL/GenBank databases. Genshu Tate, Shoin University Fujigakka Hospital, Department of Surgical Pathology, Fujigakka 1-30, Aoba-ku, Yokohama, Kanagawa 227-8501, Japan (E-mail:six10352@super.win.or.jp, Tel:81-45-974-6632, Fax:81-45-972-6242)
AUTHORS	Submitted (03-FEB-1998) to the DDBJ/EMBL/GenBank databases. Genshu Tate, Shoin University Fujigakka Hospital, Department of Surgical Pathology, Fujigakka 1-30, Aoba-ku, Yokohama, Kanagawa 227-8501, Japan (E-mail:six10352@super.win.or.jp, Tel:81-45-974-6632, Fax:81-45-972-6242)
REFERENCE	Submitted (03-FEB-1998) to the DDBJ/EMBL/GenBank databases. Genshu Tate, Shoin University Fujigakka Hospital, Department of Surgical Pathology, Fujigakka 1-30, Aoba-ku, Yokohama, Kanagawa 227-8501, Japan (E-mail:six10352@super.win.or.jp, Tel:81-45-974-6632, Fax:81-45-972-6242)
AUTHORS	Submitted (03-FEB-1998) to the DDBJ/EMBL/GenBank databases. Genshu Tate, Shoin University Fujigakka Hospital, Department of Surgical Pathology, Fujigakka 1-30, Aoba-ku, Yokohama, Kanagawa 227-8501, Japan (E-mail:six10352@super.win.or.jp, Tel:81-45-974-6632, Fax:81-45-972-6242)
REFERENCE	Submitted (03-FEB-1998) to the DDBJ/EMBL/GenBank databases. Genshu Tate, Shoin University Fujigakka Hospital, Department of Surgical Pathology, Fujigakka 1-30, Aoba-ku, Yokohama, Kanagawa 227-8501, Japan (E-mail:six10352@super.win.or.jp, Tel:81-45-974-6632, Fax:81-45-972-6242)
AUTHORS	Submitted (03-FEB-1998) to the DDBJ/EMBL/GenBank databases. Genshu Tate, Shoin University Fujigakka Hospital, Department of Surgical Pathology, Fujigakka 1-30, Aoba-ku, Yokohama, Kanagawa 227-8501, Japan (E-mail:six10352@super.win.or.jp, Tel:81-45-974-6632, Fax:81-45-972-6242)
REFERENCE	Submitted (03-FEB-1998) to the DDBJ/EMBL/GenBank databases. Genshu Tate, Shoin University Fujigakka Hospital, Department of Surgical Pathology, Fujigakka 1-30, Aoba-ku, Yokohama, Kanagawa 227-8501, Japan (E-mail:six10352@super.win.or.jp, Tel:81-45-974-6632, Fax:81-45-972-6242)
AUTHORS	Submitted (03-FEB-1998) to the DDBJ/EMBL/GenBank databases. Genshu Tate, Shoin University Fujigakka Hospital, Department of Surgical Pathology, Fujigakka 1-30, Aoba-ku, Yokohama, Kanagawa 227-8501, Japan (E-mail:six10352@super.win.or.jp, Tel:81-45-974-6632, Fax:81-45-972-6242)
REFERENCE	Submitted (03-FEB-1998) to the DDBJ/EMBL/GenBank databases. Genshu Tate, Shoin University Fujigakka Hospital, Department of Surgical Pathology, Fujigakka 1-30, Aoba-ku, Yokohama, Kanagawa 227-8501, Japan (E-mail:six10352@super.win.or.jp, Tel:81-45-974-6632, Fax:81-45-972-6242)
AUTHORS	Submitted (03-FEB-1998) to the DDBJ/EMBL/GenBank databases. Genshu Tate, Shoin University Fujigakka Hospital, Department of Surgical Pathology, Fujigakka 1-30, Aoba-ku, Yokohama, Kanagawa 227-8501, Japan (E-mail:six10352@super.win.or.jp, Tel:81-45-974-6632, Fax:81-45-972-6242)
REFERENCE	Submitted (03-FEB-1998) to the DDBJ/EMBL/GenBank databases. Genshu Tate, Shoin University Fujigakka Hospital, Department of Surgical Pathology, Fujigakka 1-30, Aoba-ku, Yokohama, Kanagawa 227-8501, Japan (E-mail:six10352@super.win.or.jp, Tel:81-45-974-6632, Fax:81-45-972-6242)
AUTHORS	Submitted (03-FEB-1998) to the DDBJ/EMBL/GenBank databases. Genshu Tate, Shoin University Fujigakka Hospital, Department of Surgical Pathology, Fujigakka 1-30, Aoba-ku, Yokohama, Kanagawa 227-8501, Japan (E-mail:six10352@super.win.or.jp, Tel:81-45-974-6632, Fax:81-45-972-6242)
REFERENCE	Submitted (03-FEB-1998) to the DDBJ/EMBL/GenBank databases. Genshu Tate, Shoin University Fujigakka Hospital, Department of Surgical Pathology, Fujigakka 1-30, Aoba-ku, Yokohama, Kanagawa 227-8501, Japan (E-mail:six10352@super.win.or.jp, Tel:81-45-974-6632, Fax:81-45-972-6242)
AUTHORS	Submitted (03-FEB-1998) to the DDBJ/EMBL/GenBank databases. Genshu Tate, Shoin University Fujigakka Hospital, Department of Surgical Pathology, Fujigakka 1-30, Aoba-ku, Yokohama, Kanagawa 227-8501, Japan (E-mail:six10352@super.win.or.jp, Tel:81-45-974-6632, Fax:81-45-972-6242)
REFERENCE	Submitted (03-FEB-1998) to the DDBJ/EMBL/GenBank databases. Genshu Tate, Shoin University Fujigakka Hospital, Department of Surgical Pathology, Fujigakka 1-30, Aoba-ku, Yokohama, Kanagawa 227-8501, Japan (E-mail:six10352@super.win.or.jp, Tel:81-45-974-6632, Fax:81-45-972-6242)
AUTHORS	Submitted (03-FEB-1998) to the DDBJ/EMBL/GenBank databases. Genshu Tate, Shoin University Fujigakka Hospital, Department of Surgical Pathology, Fujigakka 1-30, Aoba-ku, Yokohama, Kanagawa 227-8501, Japan (E-mail:six10352@super.win.or.jp, Tel:81-45-974-6632, Fax:81-45-972-6242)
REFERENCE	Submitted (03-FEB-1998) to the DDBJ/EMBL/GenBank databases. Genshu Tate, Shoin University Fujigakka Hospital, Department of Surgical Pathology, Fujigakka 1-30, Aoba-ku, Yokohama, Kanagawa 227-8501, Japan (E-mail:six10352@super.win.or.jp, Tel:81-45-974-6632, Fax:81-45-972-6242)
AUTHORS	Submitted (03-FEB-1998) to the DDBJ/EMBL/GenBank databases. Genshu Tate, Shoin University Fujigakka Hospital, Department of Surgical Pathology, Fujigakka 1-30, Aoba-ku, Yokohama, Kanagawa 227-8501, Japan (E-mail:six10352@super.win.or.jp, Tel:81-45-974-6632, Fax:81-45-972-6242)
REFERENCE	Submitted (03-FEB-1998) to the DDBJ/EMBL/GenBank databases. Genshu Tate, Shoin University Fujig

Query Match	52.6%;	Score 626;	DB 9;	Length 626;
Best Local Similarity	100.0%;	Pred. No. 1e-95;		
Matches 626;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

RRRLRFLFLILLVPAARCCGPRVGSRRPRVLVPLAKYQSPNPEKTLAS
 GRTEKTAASSERKELEPNTPNDIIFKDEENTGADLMTPORCDRLNSLAISVMQW
 PGVRLRTVEDDEHSHSESLHYEGSAVITTSRDRNRYGLIARAVAGDWMY
 ESKAHVSEKSEHSAKATGCGEPAGACVLENGEVAASAVPGRVLAEMEDGTP
 TFSVLLJEDREPRRLRAFOVETODPRRLPALILFLADHTEPAHFRATFS
 HVQGVIVSGVGLQPARVAADSTHVALSPYLPHHTEVIEDVVAECFAVADH
 HLMOLAWPLRLPFLSLMGSMTPEGVHWMPOMLYRLGRLLIBESTHPLMGAGS*

BASE COUNT 376 a 684 c 628 g 415 t
 ORIGIN

Query Match 36.1% Score 429.4; DB 12; Length 2103;
 Best Local Similarity 63.6% Pred. No. 5.3e-63;
 Matches 671; Conservative 0; Mismatches 381; Indels 3; Gaps 1;

QY 38 TGGCACTTCTGCGCGCTGCCAGCCAGAGCTCGGGCGGGCGGGCGGGCTGGCGGC 97
 DB 350 TGGTGGCTCTGGTGGCGGGCGGGCGGGCTGGCGGGCGGGCGGGCTGGCGGC 409
 QY 98 GCGCGTATGCGCGCAAGCAGCTCGTGGCGTACTCTACAGCAATTTGTCCCGCGCTGC 157
 DB 410 GGAAGCCCGCTCGCA--AGCTCGTGGCTCTTGGCTACAGCAGTTACGCCCAACGTGC 466
 QY 158 CAGAGCGGACCTGGGGCGCGAGTGGGCGAGGGAGGGGGTGGCAAGGGGCTCCGAGC 217
 DB 467 CGGAGAGACCTTGGGCGCGCGGGCGCTACGAGAGCAATGCGCGCAAGCTGTGAGC 526
 QY 218 GCTTCCGGGACCTGTCGCCCACTACACACCCGACATCTTTCAGAGATGAGAGACA 277
 DB 527 GCTTCAAGAGACTACCCCACTACATCCGACATATTTTCAAGAGAGAGAGACA 586
 QY 278 GTGAGCGGACCGCGCTGATGACCGAGCTTTCAGAGAGAGGGTGAACGCTTGGCCATG 337
 DB 587 CGGCTGGCGAGCGCTCTGACCGAGCGCTCAAGAGACCGTCTGACATCTGGCGATCT 646
 QY 338 CCGTATGAACATGTGGCGCGGAGTGGCGCTACGAGTCTGAGGGCTGGGACGAGAGC 397
 DB 647 CTGTCTATGACCAAGTGGCTGTGTGAACTCGGGGTACCGAAGGCTGGATGAGATG 706
 QY 398 GCAACCAAGCTCAGAGATTCACCTACAGAGAGCGGCTTGGACATCACTAGCTGTG 457
 DB 707 GCGATCACTAGAGAGATCTTACATATGAGGGCGGGCGGGTGAATACACACTGAG 766
 QY 458 ACCCGAGCGCAACAATATGGTGTGCGCGCGCTCGAGTGGAGAGCGGCTTGCAT 517
 DB 767 ACCGTGACCGAAATAGATGTGAGTGTGGCGCGCTTACGAGTGAAGCGCGCTTGCAT 826
 QY 518 GGGTCTACTAGACCGCGCAACAGTCCAGTGTGGTCAAAAGCTGATTAATCACTGG 577
 DB 827 GGGTGTATTAAGATTCAGAGCCCAAGTGTGTGTGTGATGATGAGATTCGGCGC 886
 QY 578 CGGTCCGGGCGGGCGGTCTTTCGGGAAATGCAACTGTGCGCTGTGAGCGCGAGC 637
 DB 887 CTGCGCAAGACAGTGGCGCTTCTTCCGCGAGCCAGGTGGCGCTGAGAGAGGGGAGC 946
 QY 638 GGAAGGCGTCCGGGAGTGCACCGCGAGACTGGGTTTGGCGCGCATCCGTACAGCC 697
 DB 947 GTGTGGCGCTTCACTGATTAAGCAGAGACCGGCTCTGCGCATGGGGAGAGATGGA 1006
 QY 698 GGGTGTGCGCCAGCGCGGTGCTTCTTCTGAGACCGGGAGTTGAGGGCGGGCTTCAT 757
 DB 1007 CCCCCACCTTATGATGATGTGCTTATTTCTTGACCGCGACCAACCGGTGAGAGCTT 1066
 QY 758 TTGTGGCTGTGAGAGCAGATGGCTTCACGCAAACTGTGCTCAGCGCTTGCACCTGG 817
 DB 1067 TCCAGGTATGAGAGATCATCTCCGCGTGGCGCTGCGCTCAGCGCTCCAGAGTGC 1126
 QY 818 TGTGTGCGCTCGAGGGCGGGCGCGCGAGGAGCTTTCACCGGTGTTCGCGGCGC 877
 DB 1127 TCTTCATTTGCGAGATATATACAGAACGAGACCGCACTTCCGCGCAATTTCCAGCC 1186
 QY 878 GGTACGCGGTGGGAGACTCGGTGTGGCGCGCGGGGAGTGGCTTGGCGAGCGGGCG 937

DB 1187 ATGTCAACCAAGCCCAATATGTGCTGTATCAGGGGTACAGCGCTCCAGCTGCTCGG 1246
 QY 938 TGGCGCGTGTGGCGCGGAGAGAACCGTGGGCGTGTCCGCGCGCTACCGCGACAGGGA 997
 DB 1247 TGGAGAGTGTCTCACCACAGAGTGGCGCTTGGGTCTATGCTCTCTACAGAGCATGGGA 1306
 QY 998 CGCTGCTGTGACAGATGTCTCTGCTTGTACGCGGTTCTTGAGAGATTCACCATGGG 1057
 DB 1307 CACTTGTGTGAGAGATGTGTGGCTCTGCTTGTGACGCTGTGACCATCATCTCG 1366
 QY 1058 CGCAGCGCGCTTGTGGCGCGCTTGTGAGTGTGCTGAC 1092
 DB 1367 CTGAGTTGGCTTGTGGCGCGCTGAGCTGTTCCTC 1401

RESULT 8
 AR021201
 LOCUS AR021201 1313 bp DNA PAT 05-DEC-1998
 DEFINITION Sequence 7 from patent US 5789543.
 ACCESSION AR021201
 VERSION AR021201.1 GI:3975816
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1313)
 AUTHORS Ingham,P.W., McMahon,A.P. and Tablin,C.J.
 TITLE Vertebrate embryonic pattern-inducing proteins and uses related thereto
 JOURNAL Patent: US 5789543-A 7 04-AUG-1998;
 FEATURES
 source Location/Qualifiers
 1..1313
 /organism="unknown"

BASE COUNT 254 a 411 c 428 g 220 t
 ORIGIN

Query Match 35.2% Score 419.8; DB 5; Length 1313;
 Best Local Similarity 64.5% Pred. No. 2.3e-61;
 Matches 701; Conservative 0; Mismatches 362; Indels 24; Gaps 4;

QY 41 CACTTCTGGGGCTGCGCAGCCAGAGTGTGGCGGGCGGGCGGGCGGGCTGGCGGC 100
 DB 47 CGGTGTGTGTGTCGCCCGGGCGTGTGGCGGGCGGGCGGGCGGGCTGGAGAGCGGC 106
 QY 101 GCTATGGCGCAGCAAGCAGTGTGCTGCTTCTTACAAAGCAATTTGTGCGCGGTGCGAG 180
 DB 107 AC-----CCCAAAAGCTACACCCCTTTTACCTACAGAGAGCTTATTTCCAGCTAGCG 180
 QY 161 AGCGGACCTGGGCGCGCAGTGGGCGAGCGGAGGGAGGGGTGGCAAGGGCTCCGAGCGCT 220
 DB 161 AGAAGACCTTATGAGGGCGCGAGCGAGATATGAGGAGAAAGATCAACAAGAACTCCGAGCAT 220
 QY 221 TCCGGGACCTGTGGCCCAATACAGCCCGCAATCTTCAAGATGAGAGAAACAGT 280
 DB 221 TTAAAGAACTACCCCAATTAACACCCGAGACATCATATTAAGATGAGAGAAACAGC 280
 QY 281 GAGCGGACCTGTATGACCGAGCTTGCAGAGAGAGAGGTGAAGCGCTTGGCCATTGGCG 340
 DB 281 GAGCAGACCGGCTGTATGACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 340
 QY 341 TGATGAACATGTGGCGCGGAGTGGCGCTTACAGTGAAGTGAAGGCTGGAGAGAGAGAG 400
 DB 341 TGATGAACATGTGGCGCGGAGTGGCGCTTACAGTGAAGTGAAGGCTGGAGAGAGAGAG 400
 QY 401 ACCAGCGTCAAGATTAATCACTACAGAGCGGCTTGTGAGACATCACTAGCTGTGAGC 460
 DB 401 ATCATTTAGAGAGATCTTACATATGAGGTTGAGACAGTGAAGTGAAGTGAAGTGAAG 460
 QY 461 GCAACGCAACAGATATGAGTGTGCTGGCGCGCTTCAGTGAAGCGCGGCTTGGATGGG 520
 DB 461 GGAACCGCAGACAGTACGAGTGTGCTGGCGCGCTTGGCTGTGGAAGAGAGTGTGACTGG 520

QY	521	TCTACTAGAGAGCCCGCAACACAGTCGACGTCGTGGGTGTAAGAGCTATACACACGCGGG	580
Db	521	TCTACTAGAGATCCAAAGCTACATCCATCTTCTGTAAAGCAGAGACTCCGTGGCG	580
QY	581	TCCGGGCGGGGCGGTGCTTTCCGGGAAATGCAACATGTGCGCCTGTGTGAGCGCGCAGCGGA	640
Db	581	CCAAATCCGGGCGGTGTTCCTCCGGGATCCGCAACCGTGCACCTGGAGAGGCGCGCACCA	640
QY	641	AAGGCTCCGGGAACTGCACACGCGGAGACTGGGTTTGGCGGCGCATGCTGAGCGCG	700
Db	641	AGCTGTGTAAGACTTACGTCCCGGAGACCGGTGCTGTGGCGCTACGACACAGGCGCGC	700
QY	701	TGGGCCACACGCGGTGCTGCTCTTCCTGAGACCGGAGACTTGCAGCGCGGCTTATTG	760
Db	701	TGCTGTACAGGACTCTCTACACTTCTCTGGACCGGACGAGAGGCGCCAAAGAGTCTTCT	760
QY	761	TGGCTGTGAGACCGAGTGGCTCCACGCAAACTGTGCTCACGCGCTGGACCTGTGT	820
Db	761	ACGTGATGAGACGCTGTGAGCGCGCGAGACGCGCTGTCTGTACCGCGGACCTGTCT	820
QY	821	TTGCGCGTCT-----GAGGCGCGCGCCCGCGCCACGCGACTTTGCACCGGTGT	868
Db	821	TGTGGGCCCGCACACACTCGGGGCGCCACGCGCGCGCGCAAGCGCGCTTTTGCAGCC	880
QY	869	TGCGCGCGCGGTACGCGCTGGGAGACTGGTGTGTGGCGCGCGCGCGGGA---TGCGCTTC	925
Db	861	CGGTGCGCGCGCGGTGAGCGCGCTGTGCTGTGAGACGCGCGCGCGCGCGCGCGGTGC	940
QY	926	GGCCAGCGCGCGGTGGCGCGGTGTGGC---GCGGAGAGAAAGCCGTGGCGGTGTGCGCGCG	982
Db	941	TGCCGCGCGCGGTGCACAGCGTGAAGCGTGCAGAGAGAGAGGCGGGCGGTACGCGCGCG	1000
QY	983	TCACCGCGCGAGGGAACGCTGCTGGTGAACGATGTCTCTGGCTCTTGTCTACGCGGTTCG	1042
Db	1001	TCACGCGCGCGCGCACCACTTCATCAACACCGGCTGCTGCGCTGCTGCTGCTGCTGCTG	1060
QY	1043	AGAGTACACAGTGGGCGGACCGCGCTTTTGTCCCTTGTGAGACTGCTGCACGCGGTACGG	1102
Db	1061	AGGACACAGCTGGGCGACACACGCGCGCTTCGCGCTTTCCGCTTGGCGCGCACGCGCTGTG	1120
QY	1103	CGCTGCT 1109	
Db	1121	CCGCGCT 1127	
RESULT	9		
LOCUS	AR063083	1313 bp	DNA
DEFINITION	Sequence 4 from patent US 5844079.	PAT	29-SEP-1999
ACCESSION	AR063083		
VERSION	AR063083.1	GI:5990774	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 1313)		
AUTHORS	Ingham,P.W., McMahon,A.P. and Tabin,C.J.		
TITLE	Vertebrate embryonic pattern-inducing proteins, and uses related thereto		
JOURNAL	Patent: US 5844079-A 4 01-DEC-1998;		
FEATUES	Location/Qualifiers		
SOURCE	1..1313		
BASE COUNT	254 a	411 c	428 g
ORIGIN	220 t		
	/organism="unknown"		

	Query Match	35.2%	Score 419.8;	DB 5;	Length 1313;
	Best Local Similarity	64.5%;	Pred. No. 2.3e-61;		
	Matches 701; Conservative	0;	Mismatches 362;	Indels 24;	Gaps
07	41 CACTTCTGGCCCTGGCCACCCAGAGACTCGGGGGCCGGGGGGCCGGTGTGGCCGGGCC	100			

D	b	47	CGCTGCTGCTGTGCTGCCCGGGCTGAGCTGTGAGGCCGCGACGAGGGGCTTTGGAAAGAGGGGC	106
Q	y	101	GCATGCGCGCAAGCAAGCTGTGTCGGCTACTCTTACAGCAATTTGTGCCGCGCTGCCAG	160
D	b	107	AC-----CCCAAAAGACTGACCCCTTTTAGCTTACAGCAAGTTTATTCCCAAGCTAGCCG	160
Q	y	161	AGCGGACCCCTGGGGCGCAGTGGGGCCAGGGGAGGGGAGGGGTGGCAAGGGGGCTCCGAGGCT	220
D	b	161	AGAGACCCTTAGGGGCCAGGGCGAGATATGAAGGGAGATCAAGAAAGATCCGAAGCAT	220
Q	y	221	TCCGGAGCTGCTGGCCCAATACACCCGACATCATCTTCAGGATAGAGGAGAAGTAGT	280
D	b	221	TTAAGGAAGCTACCCCCCAATTAACACCCGACATCATATTTAAGATAGGAAAGAACGG	280
Q	y	281	GAGCGGACCCGCTGATATGACCGAGCGTTGGACAGAGAGGGGTGAACCTTTGGCCATTGCG	340
D	b	281	GAGCAGACCCGCTGATGACTCATAGAGGTGCAGAAAGACAAAGTTAATCCTTGGCCATCTG	340
Q	y	341	TGATGAACATGTGGCCCGGAGTGGCTTACGATGACTAGGGCTGGGACAGAGACGGCC	400
D	b	341	TGATGAACACAGTGGCCCGGAGTGAAGGCTCGAGATGCCAGGGGCTGGGATGAGACGGCC	400
Q	y	401	ACCGAGCTCAGGATTCACCTACATAGAGAGCGCGTCTTTGACATCACTACGTGTGAC	460
D	b	401	ATCATTTCAAGAGGATCTCTACACTATGAGGGTGTGACAGTGAACATCAACAGTCCGACC	460
Q	y	461	GCGACCGCAACAAATATGGGTTGCTGGCGCGCCTCGCATGTGAACCCGAGCTTCGACTGG	520
D	b	461	GGGACCGGAGCAATAGCGGATCGTGGCTGGCTGGCTGGAGACAGATTTCCAGTGG	520
Q	y	521	TCTACTACGAGTCCCCGCAACCAACGCTCAGGTGTGGTCAAAGCTGATATCATCTGGCGG	580
D	b	521	TCTACTATGAATCCAAAGCTCAATCACTACGTCTGTGTGAACCAAGAACTCCGTGGCG	580
Q	y	581	TCCGGGGGGGGGTGGCTTTCGCGGAAATGCAACGTGTGGCGCTGTGAGAGCGCGAGCGGA	640
D	b	581	CCAAATCGGCGGCTGTTTCCCGGGATCCGCAACCTGTGAGCAAGGAGCGGACCA	640
Q	y	641	AAGGACTGCGGGAACCTGACACCGCGGAGACTGGGTTTGGCGGCGCATGGCTAGGCCGGG	700
D	b	641	AGCTGTGAAGGACTTACGTCCCGGAGACCGGCTGCTGGCGGCTGACGACAGGCGCGG	700
Q	y	701	TGCTGCCACGCGCGGTGCTCTTCTCTGAGACCGGAGACTTGCAGCGCGCGGCTTATTG	760
D	b	701	TGCTGTACAGGACTTCTCATCTTCTGTGACCGGACGACGACGAAGGCGCAAGAGTCTTCT	760
Q	y	761	TGGGTGAGGAACCGGAGTGGGCTTCAGAGCAAAATGTGTGTACGCGCTGGACCTGGGT	820
D	b	761	ACGTGATGAGACCTGTGAGACCCCGCGAGCGCGCTGTGTCTACACCGCGGCACTGTGCT	820
Q	y	821	TTGCGCGT-----GAGGGCGGGCGCCGCGCAGAGCGACTTGGACCGGTGT	868
D	b	821	TCTGTGGCCCGCCAAAGAGACTCGGGGCCACGCGCGGCGCAAGCGCGCTTTTGGCACCC	880
Q	y	869	TGCGCGCGCGGCTACGCGCTGGGGGACTCGGTGCTGGCGCGCGCGGGGA--TGCCTTC	925
D	b	881	GGCTGGCGCCCGGCGCAGCGGTGAGCTGAGTGGTGTGCTGAACGCGCGGGGACCGCGGCTGC	940
Q	y	926	GCCCAAGCGCGGTGGCCCGGTGTGGC---GGGGAGGAACCTGGGGCGTTCGGCGCGC	982
D	b	941	TGCGCGCGCGGCTACAGAGGTGACGCTGGGAAGGAGGAGCGGGCGGCTACGGCGCGC	1000
Q	y	983	TCACGCGCAGGGAGCGCTGTGGTGAAGCATTCTTGCGCTTGTGCTACGAGGGGTTCGG	1042
D	b	1001	TCACGCGCAGGGAGCAATTCATCATCAACCGGGTGTCCGCTCTGTGCTACGCTGTCAATCG	1060
Q	y	1043	AGAGTCAACAAGTGGCGCACCGCGCTTTTGGCCCTTGAAGCTGTGCAACGCGCTAGGGG	1102
D	b	1061	AGGAGCAACAGCTGGGACACACCGGGCGCTTGGCGCTTTCGGCTGGGGGACACGCGGTGCTGG	1120
Q	y	1103	CGCTGCT 1109	
D	b	1121	CGCGGCT 1127	

	DB	1061	AGGACACAGCTGGGACACC GGCGCCTTGCCGCTGGCGGCAGCGCGTGTGG	1120
	QY	1103	CGCTGCT	1109
	DB	1121	C G C G C G C T	1127

REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	FEATURES
1 (bases 1 to 1715)	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. Roelink H., Augsburg A., Heemsterk J., Korzh V., Norlin S., Ruiz i Altaba A., Tanabe Y., Placzek M., Edlund T., Jessell T.M. and et al.	Floor plate and motor neuron induction by vhh-1, a vertebrate homolog of hedgehog expressed by the notochord	Cell 76 (4), 761-775 (1994)	94170375	Location/Qualifiers 1..1715 /organism="Rattus norvegicus" /strain="Sprague-Dawley" /db_xref="taxon:10116" 315..377 /gene="vhh-1" 315..1628 /gene="vhh-1" 315..1628 /gene="vhh-1" /standard_name="vertebrate homolog of hedgehog" /codon_start=1 /protein_id="AA20999.1" /db_xref="gi:452123" /translation="MLLLARCLFLVALASLLVCPGLAGCGRGKROHPKLTPLPAVKRISPNVAKETLGASGRGKTRNSRFEPLRPNNPDITFDEPTGADRLMTRCKDLMALAIIVNMQMPVKRLRTLEGDEDDHSESLHTGKRAVDITTSRDSKTKGLALVAEAGFDWVYTESKARIHCSVAENSVAAKSDCEFGSAFVHLEDGTLVNDLSPDRYLAADDQGRILYSDLEPLDRBEGKVKVYFETEPERERLLTPAHLFPVPHNSGPTPGSPPLFASVRPGRGVYVAEGRGRYLLPAVHSVTTLREAGAAVPLPADDTILINRYLASCYAVIEHSMWNRFAFRLAHLALALAAFAARDGGGSGSIPAPQSVAEARGACGPAGIHHWYSQLYHICGTMWLDSETHPLGMAVKSS"
BASE COUNT	355 a	518 c	567 g	275 t	
ORIGIN					
Query Match	34.2%	Score 407;	DB 12;	Length 1715;	
Best Local Similarity	63.8%	Pred. No. 3.1e-59;			
Matches 693;	Conservative 0;	Mismatches 370;	Indels 24;	Gaps 4;	
41	CACCTTGTGGCGCTGCCAGCCAGACCTGGCGGGCGGGCGCGGTGGCCGCGCC	100			
Db	361 CCGCTCGTGTGAGCCCGGACGTGGCTGTGGCGGGCGGGCGGTGGGAAGGACAGC	420			
QY	101 GCTATGGCGGCGAGAGCTCGTGGCGCTACTTACAGCAATTTGTGCCGGCGTCCAG	160			
Db	421 AC-----CCCAAAAAGCTGACCCCTTTAGCTTACAGCAGATTATCCCACTAGCCG	474			
QY	161 AGCGAGCCTGGGGCGCGAGTGGCGGACGAGGGAGGGAGGGTGGCAAGGGCTCCGACCT	220			
Db	475 AGAAGACCCCTTAGGGCGCGCGCGCGATTTGAAGGGAAGTACAAAGAACTCCGAACAT	534			
QY	221 TCCGGAGCCTGCTGGCCACTACAAACCCCGCATCATCTTCAAGAGATGAGAAACAGTG	280			
Db	535 TTAAGGAGCTACACCCCAATTACAAACCCCGCATCATTTTAAGATGAGAAAACATG	594			
QY	281 GAGCGGACCGGCTGATGACGAGCGCTTCAAGAGAGAGGGTGAACGCTTTGGCCATTGGC	340			
Db	595 GAGCAGACCGGCTGATGACTCAGAGGTGCAAAAGCAAGTTAAATGCTTTGGCCATCTCG	654			
QY	341 TGATTAACATGTGGCGCGCGGAGTGGCCCTTACGAGTACATGAGGGCTGGGACGAGACGGC	400			
Db	655 TGATTAACCATGTGGCGCGGAGTGGAGGCTTGGAGTACGAGGGCTGGGATGAGACGGCC	714			
QY	401 ACCAGCGTACAGATTCACATCCACGAAAGGCGGTGGTGGACATCACTAGCTGAGCC	460			
Db	715 ATCATTCAGAGAGCTCTTACACTATGAGAGGTGAGACAGTGGACATCAACAGTGTGACA	774			
QY	461 GCGAGCGCAACAGATGGGTGCTGGGGCGGCGCTGCAGTGGAAACCGGCTCGACTGG	520			
Db	775 GGGAGCGGAGCAAGTATGCAATGCGTGGCGCTGCGCTGTGAGAGCTGGATTGCACTGG	834			
QY	521 TCTACTACGAGTCCCGCAACACGATCCAGTGTGGGTGAAGCTGATTAATCACTATCGCGG	580			

[illegible]

CDS

/db_xref="taxon:8355"
1..1191
/note="secreted signaling molecule; Method: conceptual translation supplied by author."
/codon_start=1
/product="cephalic hedgehog protein"
/protein_id="AA85163.1"
/db_xref="GI:1147828"
/translation="MPAVRIYVILAIICGLLIVRCGPGRPVRRRNRKLYPLHY
KOFVNPETKLGASRSEKIHRSERFELVNPYDIIKDEKRGADRLTERC
KDRYNALAIYVMNMWPGVKLRHEGMEDEPHADSLHYGRALDITSDRNRKYM
LARLAVAGFDWYTESKAHIVSYKADNSIGRASGCCFPGTANMAGTERPLSEL
RIGDTVTIDETGLITSVLELHRNPYATFVLIEAGHPSKLVLPNHLFIOS
SSSAGFLPTFAIVQIDLVQIYVNGQVOSKRVVRSLEQTEVAPMTEHGLTLD
GVLTSCATVESHIAHVSIAPLRLFOGIASMLPDLMSGVHMYCHILVLAAYVLM
MDMP"

BASE COUNT 299 a 274 c 312 g 306 t
ORIGIN

Query Match 34.1%; Score 406.6; DB 4; Length 1191;
Best Local Similarity 61.9%; Pred. No. 3.8e-59;
Matches 719; Conservative 0; Mismatches 424; Indels 18; Gaps 4;

QY 33 CTGCTGGCACTTCTGCGCTGCGCCAGCCAGAGCTGCGGGCCGGGCGCGGCTTGG 92
DB 33 CTGCTGGCGGCTCTCCCTGCTCCAGTTCGCTGTGTGTGCTGCGCGGCGGCTTGG 92
QY 93 CCGGGCGGCGTATGCGGCGCAAGAGCTGCGGCTGCTACTACAGCAATTTGTCGCCG 152
DB 93 CCGTGGCGCTTACATAGAGAAA--CTAGTCCCACTGCACTACAGAGTTGCTCTTAA 149
QY 153 CGTCCAGAGCGGACCCCTGGGCGCCAGTGGCCAGCGAGGGAGGTGGCAAGGGGCTC 212
DB 150 TGTGGCAAGAAAGACTAGAGGGCCAGTGGCAAGTCAAGGCAAGATTCACCGGGGCTC 209
QY 213 CGAGGCGTTCGGGGA--CTGTGCCCCACTACAAACCCGACATCTTCAAGAGATGAGA 272
DB 210 AGAAAGTTCACTGAGTGTGCTCCCACTACAAACCAACATCATCTTCAAGAGATGAGA 269
QY 273 GAACAGTGGACCCGACCCCTGATGATACCGAGCTTGCAGAGAGAGGTGAACGCTTGGC 332
DB 270 AAAAAGAGAGCTAGACAGCTCATGATGACGCGCTGTAAAGACAGATCAATGCTTGGC 329
QY 333 CATTGCGCTGATGAACATGATGAGCGGAGTGGCGCTTACAGTGAAGTGGGCTGGAGCA 392
DB 330 TATCTCTGTGATGAACATGTGGCCAGGTGTGAAGCTAGCGGTACCGAGGTGGAGAGA 389
QY 393 GGAGCGCCACACGCTCAGATTCACTCAGAGAGCGGTGCTTGGACATCACTAC 452
DB 390 GGATGGACACACGCTCAGACTCTCATATATGAAAGGCGCGCACATGACATTAACAAC 449
QY 453 GTCGACCGCGACCCGCAACAGTATGGGTGTGCGGGCGCTCGAGATGGAAACCGGCTT 512
DB 450 ATCTGACCGAGACAGAAATAGGCAATGTGGACGTGAGAGTGAAGCTGGCTT 509
QY 513 CGACTGGGTACTAGAGTCCCGCAACAGCTGATGCGGTGCTCAAGTGAATCACTC 572
DB 510 TGACTGGGTGTATGATCAAGACACATCATGTCTGTGMAAGCAGATTAATC 569
QY 573 ACTGGCGTCCGGGCGGCGGCTGCTTCCGGGAAATGCAACTGTGCGGCTGTGAGCGG 632
DB 570 CCGTGGGTGCGGCTGTGAGGCTGTTTCAAGAACAGTATGATATGATGGAGACTGG 629
QY 633 CGAGCGGAAAGGGCTGCGGGAAGTGCACCGGAGACTGGGTTTGGGGCGCGATGCGTC 692
DB 630 TGAGGAAACCACTGTGTGAAGTAAAGTGTGATGCTCTCACTACATCACTGAAC 689
QY 693 AGGCGGGGTGAGCCACGCGGCTGTCTTCTGAGCGGGAGCTTGCAGCGCGCGGCG 752
DB 690 AGGCGAGTTAATTAATTAAGTGTCTGTCTTCTTGCACAGGAACCCATATAAAGAAC 749
QY 753 TTCATTTGTGGCTGTGAGAGACGAGTGGCGTCCAGCAAACTGTTGCTACGCGCTGGCA 812

DB 750 TACGTTGTGCTTATTTAGCTGAGGAGACACCCATCCAGCTCTTGTACCCCAATCA 809
QY 813 CCGTGTGTTGCGCGCTGAGGGCGGGCGCGCCAGAGGAGACTTTCACCGGCTTCGC 872
DB 810 TCTGTGTTTATTAACAAACAG-----CTCATGACAGAGATTTCTACCACTTTTGC 860
QY 873 GCGCGGCTTACGCGCTGGGAGACTGCTGCTC--TGCGCCCGCGCGGATGCGCTTTCGCC 929
DB 861 TTATAGGTTAGATAGAGAGACCTGTTCAAAATATATGTAATGCACTGAGTGCATC 920
QY 930 AGCGCGGCTGCGCGCTGTGCGCGGAGAAAGCCGTGGCGGTTCGCGCTCTACCCC 989
DB 921 TTTCAAGTTGTGAGGGATCTCTTAGAGAAACAACTGGAATATATCTCTATAGTGA 980
QY 990 GCAGCGGCGCTGCTGGAGAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1049
DB 981 ACATGAAACCTTACTTGTGATGGGCTGCTACATCTCTATCTCTACTGTGAGTCCCA 1040
QY 1050 CCACTGGGCGCAGCGCGCTTTTGGCCCCCTTGAAGTCTGTCAGCGCTAGGCGGCTGCT 1109
DB 1041 CACATTTGSCACATGATATCTTGGCCCCACCTTAGACTCTTCAAGCAATAGCTCAATGCT 1100
QY 1110 CCGCGGCGGCGCGCTCCAGCCGACTGCGATGCTGCTGCTGCTGCTGCTGCTGCT 1169
DB 1101 CCCGATTTTACACAT--GAGTGAATGGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1157
QY 1170 AGCGAGGAGCTACTGCGGCTG 1190
DB 1158 GCCAAGTATGTTCTTGTGCT 1178

RESULT 14
AR063085
LOCUS AR063085 1425 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 6 from patent US 5844079.
ACCESSION AR063085
VERSION AR063085.1 GI:5990776
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1425)
AUTHORS Ingham, P.M., McMahon, A.P. and Tabin, C.J.
TITLE Vertebrate embryonic pattern-inducing proteins, and uses related thereto
JOURNAL Patent: US 5844079-A 6 01-DEC-1998;
FEATURES
source Location/Qualifiers
1..1425
BASE COUNT 249 a 461 c 506 g 206 t 3 others
ORIGIN

Query Match 32.8%; Score 390.6; DB 5; Length 1425;
Best Local Similarity 61.9%; Pred. No. 1.7e-56;
Matches 711; Conservative 0; Mismatches 369; Indels 69; Gaps 3;

QY 18 TCTACTGCTTGTGCTGCTTGGCACTTCTGCGCTGCGACCCAGAGACTGCGGCGG 77
DB 21 TCTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 80
QY 78 CCGGGGCGCGGTGCGGCGCGCGCTGAGCGCGCAAGACGTCGTCGCTACTTCAAA 137
DB 81 CAGGGGCTTCGGGAAGAGAGAGC-----CCAAAAAGCTGACCCCTTTAGCTACAA 134
QY 138 GCAATTTGTGCGCGGCTGCGAGCGGACCTTGGCGCCAGTGGCGCCAGCGGAGGAG 197
DB 135 GCAATTTTCCCAATGATGCGGAGAAAGACCTTAGCGGCGCAGGAGATGAAGGAA 194
QY 198 GGTGGCAAGGGGCTTCGAGCGCTTCGCGGACCTGTGCGCACTACAAACCCGACATCAT 257
DB 195 GATCTCCAGAAACCTCCGAGCATTTAAGAACTCACCCCAATTAACAACCCGACATCAT 254


```
Db 466 GTTGAACGCTTTGGCCATTCGGTGTATGAACAGTGGCCAGAGTGAATGCGGGTGAC 525
QY 378 TGAGGGCTGGGACGAGAGCGCCACCACGCTGAGATTCTACATACGAAGCCGTCG 437
Db 526 CGAGGGCTGGGACGAGAGTATGGCCACACTCAGAGAGTCTCTGCACTACGAGGGCCGCGC 585
QY 438 TTTGACATCACTACGTCTGACCGCGACCGCAACAATATGGGTGCTGGCCGCTCGC 497
Db 586 AGTGGACATCACACGTCTGACCGCGACCGCAAGTACGGCATGTGGCCGCTGGC 645
QY 498 AGTGAAGCCGGCTTGAGTGGGTACTACGATCCCGCAACCACTGACGTGCGGT 557
Db 646 GGTGGAGCGCGGCTTGAGTGGGTACTACGATCCAGGACATATCCACTGCTGGT 705
QY 558 CAAAGCTGATTAATCACTAGCGGCTCCGGCGGGCGGCTGCTTCCGGAAATGCAACTGT 617
Db 706 GAAAGCAGAGAACTCGGTGGCGCAATCGGAGGCTGCTTCCGGGCTCGGCGACAGGT 765
QY 618 GCGCTGTGGAGCGGCGGAGCGGAAAGGCTGCGGAACTGACCGCGAGACTGGGTTT 677
Db 766 GCACCTGGAGCAGGGGCGACCAAGCTGGTGAAGGACTGAGCCCGGGGACCGCTGCT 825
QY 678 GGGGCGCGATGGGTGAGGGCGGGTGGTGGCCAGCGCGGTGCTCTCTCTGAGCCGGGA 737
Db 826 GGGGCGGAGCAGACAGGCGCGGCTGCTCTACAGGACTTCTCTCTCTGAGCCGGA 885
QY 738 CTTGACGCGCGGCGCTTCACTTGTGCTGTGAGACCGAGTGGCTTCAAGCAACTGTT 797
Db 886 CGAGCGCGCGCAAGAGTCTTCTAGCTGATCGAGAGCGGAGCCCGCGAGCGCTGCT 945
QY 798 GCTCACGCGCTGGCACCCTGCTGTTGCCG-----TCGAGGCGCGCGCGCGCG 846
Db 946 GCTCACGCGCGCGCCTCTCTTGTGGCGCGCGCACAAAGACTCGGCCACCGGGAGCC 1005
QY 847 CCAGGGGACTTTGCAACGGTGTTCGGCGCGCGGCTACGCGCTGGGAGACGCTGCTGGCG 906
Db 1006 CGAGGCGTCTCGGGCTCGGGCGCGCTTCGGGGGCGACTGGGCGCTGGGCGCTGTT 1065
QY 907 CCGGCGGGGATGCGCTTTCGGCCAGCGCGGTGGCCGTG-----947
Db 1066 CGGCAAGCGCGGTGCGCGCGCGCGGCGGCTGCTGCTGCGGAGCGTACCGGGAGCG 1125
QY 947 -----TGGCGGGGAGAAACGCTGGCGTGT 974
Db 1126 CGGGCTCTGCGCGCGCTGTGCAAGCGTGAACCTTAAGCGAGAGGCGCGGGCGCTTA 1185
QY 975 CCGGCGCGCTACCGCGCGGAGCGGAGCGTGTGTAAGATGTCTGCGCTCTGCTACGC 1034
Db 1186 CGGCGCGCTACCGCGCGGAGCGGAGCGATTCATCAACCGGGTGTGCGCTCTGCTACGC 1245
QY 1035 GGTCTGAGAGTCAACAGTGGGCGGACCGCGCTTTGCGCCCTTGAGACTGCTGCACGC 1094
Db 1246 GGTCAATCAGAGAGCAAGCTGGGCGGACCGGGCTTTCGCGCTTCCGCTGGCGCACGC 1305
QY 1095 GCTAGGGGC 1103
Db 1306 GCTCTGGC 1314
```

Search completed: June 5, 2000, 07:33:10
Job time: 9655 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2000, 06:53:49 ; Search time 52.13 Seconds
(without alignments)
4044.423 Million cell updates/sec

Title: US-08-900-220-7

Perfect score: 1622
1 CATCACCCACGAGAGACC.....CATTTGGAGGCCCATTCCTCC 1622

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

Issued Patents: NA: *
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/ina/5C.COMB.seq: *
4: /cgn2_6/ptodata/1/ina/5D.COMB.seq: *
5: /cgn2_6/ptodata/1/ina/6.COMB.seq: *
6: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq: *
7: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	936	57.7	939	3	US-08-356-060A-7 Sequence 7, Appli
2	843.2	52.0	1056	2	US-08-176-427B-5 Sequence 5, Appli
3	843.2	52.0	1056	3	US-08-356-060A-3 Sequence 3, Appli
4	459	28.3	1277	2	US-08-176-427B-1 Sequence 1, Appli
5	459	28.3	1277	3	US-08-356-060A-1 Sequence 1, Appli
6	453.8	28.0	1425	3	US-08-356-060A-6 Sequence 6, Appli
7	452.2	27.9	1576	2	US-08-748-591-5 Sequence 5, Appli
8	452.2	27.9	1576	2	US-08-748-591-10 Sequence 10, Appli
9	438.8	27.1	1313	2	US-08-176-427B-7 Sequence 7, Appli
10	438.8	27.1	1313	3	US-08-356-060A-4 Sequence 4, Appli
11	432	26.6	1190	6	PCT-US95-02315-1 Sequence 1, Appli
12	419.8	25.9	1190	2	US-08-176-427B-3 Sequence 3, Appli
13	419.8	25.9	1190	3	US-08-356-060A-2 Sequence 2, Appli
14	418.2	25.8	1256	2	US-08-176-427B-9 Sequence 9, Appli
15	418.2	25.8	1256	2	US-08-356-060A-5 Sequence 5, Appli
16	144	8.9	144	6	PCT-US95-15463-2 Sequence 2, Appli
17	144	8.9	144	6	PCT-US95-15463-1 Sequence 1, Appli
18	112	6.9	144	6	PCT-US95-15463-1 Sequence 1, Appli
19	112	6.9	144	6	PCT-US95-15463-1 Sequence 1, Appli
20	47.6	2.9	2887	5	US-09-183-253-1 Sequence 1, Appli
21	47	2.9	44377	4	US-08-804-227C-7 Sequence 7, Appli
22	47	2.9	44377	4	US-08-804-198-1 Sequence 1, Appli
23	44.8	2.8	1433	4	US-08-666-392A-1 Sequence 1, Appli
24	43.8	2.7	765	5	US-08-718-904-79 Sequence 79, Appli
25	43	2.7	4524	2	US-08-845-998-7 Sequence 7, Appli
26	42.8	2.6	2004	2	US-08-471-033-18 Sequence 18, Appli
27	42.8	2.6	2004	3	US-08-471-044-18 Sequence 18, Appli

28	42.8	2.6	2004	3	US-08-463-483A-18 Sequence 18, Appli
29	42.8	2.6	2004	3	US-08-471-046A-18 Sequence 18, Appli
30	42.8	2.6	2004	3	US-08-470-566B-18 Sequence 18, Appli
31	42.8	2.6	2004	4	US-08-469-334-18 Sequence 35, Appli
32	42.8	2.6	2576	2	US-08-471-033-35 Sequence 35, Appli
33	42.8	2.6	2576	3	US-08-471-044-35 Sequence 35, Appli
34	42.8	2.6	2576	3	US-08-463-483A-35 Sequence 35, Appli
35	42.8	2.6	2576	3	US-08-470-566B-35 Sequence 35, Appli
36	42.8	2.6	2576	4	US-08-469-334-35 Sequence 35, Appli
37	42.8	2.6	2576	4	US-08-471-033-17 Sequence 17, Appli
38	42.8	2.6	2655	2	US-08-471-033-26 Sequence 26, Appli
39	42.8	2.6	2655	3	US-08-471-044-17 Sequence 17, Appli
40	42.8	2.6	2655	3	US-08-471-044-26 Sequence 26, Appli
41	42.8	2.6	2655	3	US-08-463-483A-17 Sequence 17, Appli
42	42.8	2.6	2655	3	US-08-463-483A-26 Sequence 26, Appli
43	42.8	2.6	2655	3	US-08-471-046A-17 Sequence 17, Appli
44	42.8	2.6	2655	3	US-08-471-046A-26 Sequence 26, Appli
45	42.8	2.6	2655	3	US-08-471-046A-26 Sequence 26, Appli

ALIGNMENTS

RESULT 1
US-08-356-060A-7
Sequence 7, Application US/08356060A
Patent No. 5844079
GENERAL INFORMATION:
APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tablin, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
Proteins and Uses Related Thereto
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHYE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356, 060A
FILING DATE: 14-DEC-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/176, 427
FILING DATE: 30-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMT-006CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 939 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..939
US-08-356-060A-7

Query Match 57.7%; Score 936; DB 3; Length 939;
 Best Local Similarity 100.0%; Pred. No. 1.7e-198;
 Matches 936; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 351 CGCCTATGACCCAGCGCTGACGAGGAGCGGCTGAACTCGCTGCTATCTGGTGTGATGAAAC 410
DB 4 CGCCTATGACCCAGCGCTGACGAGGAGCGGCTGAACTCGCTGCTATCTGGTGTGATGAAAC 63
QY 411 CAGTGGCCCGGTGTGAAGCTGCGGGGTGACCGAGGAGGTGGAGACGAGGAGCCACTGCA 470
DB 64 CAGTGGCCCGGTGTGAAGCTGCGGGGTGACCGAGGAGGTGGAGACGAGGAGCCACTGCA 123
QY 471 GAGGAGTCCCTGCATTATGAGAGGCGCGCGGTGACATCACACATACAGACCGGAGCCG 530
DB 124 GAGGAGTCCCTGCATTATGAGAGGCGCGCGGTGACATCACACATACAGACCGGAGCCG 183
QY 531 AATAGATGAGAGTGCCTGGGCGGCTTGACAGTGGAGGCGGCTTGACTGGGTATATAC 590
DB 184 AATAGATGAGAGTGCCTGGGCGGCTTGACAGTGGAGGCGGCTTGACTGGGTATATAC 243
QY 591 GAGTCAAGGCGCCAGCTGCAATGCTCCGTCAAGTCCGAGCACTCGCCGCAAGCCAGAGACG 650
DB 244 GAGTCAAGGCGCCAGCTGCAATGCTCCGTCAAGTCCGAGCACTCGCCGCAAGCCAGAGACG 303
QY 651 GCGGCGTCTTCCCTGCGGAGCCAGAGTACGCTGAGAGTGGGGCGGCTGTGGCTTG 710
DB 304 GCGGCGTCTTCCCTGCGGAGCCAGAGTACGCTGAGAGTGGGGCGGCTGTGGCTTG 363
QY 711 TCAGCGCTGAGCGCGGAGACCGGTGTGCTGGCCATGGGGGAGATGGAGACCCCACTTC 770
DB 364 TCAGCGCTGAGCGCGGAGACCGGTGTGCTGGCCATGGGGGAGATGGAGACCCCACTTC 423
QY 771 AGCGATGCTCATTTCTCTGAGACCGGAGCCAGCCAGGCTGAGAGGCTTCCAGGTATC 830
DB 424 AGCGATGCTCATTTCTCTGAGACCGGAGCCAGCCAGGCTGAGAGGCTTCCAGGTATC 483
QY 831 GAGACTAGAGACCCCGACGCGGCTGCACTACACCCGCTACCTGCTTTAGCGCT 890
DB 484 GAGACTAGAGACCCCGACGCGGCTGCACTACACCCGCTACCTGCTTTAGCGCT 543
QY 891 GACATACACGAGGAGCGGAGCGGCTTCCGGGCAATTTGCGAGCAAGTGGACGCT 950
DB 544 GACATACACGAGGAGCGGAGCGGCTTCCGGGCAATTTGCGAGCAAGTGGACGCT 603
QY 951 GCGCAGTACGTGTGTGCTGAGGAGTCCAGGCTGCAAGCTGCGGCTGCGAGCTGTC 1010
DB 604 GCGCAGTACGTGTGTGAGGAGTCCAGGCTGCAAGCTGCGGCTGCGAGCTGTC 663
QY 1011 TCTACACAGTGGCGCTGCGGAGCTACGCGCGCTCACAAAGCATGGGACATGGTGTG 1070
DB 664 TCTACACAGTGGCGCTGCGGAGCTACGCGCGCTCACAAAGCATGGGACATGGTGTG 723
QY 1071 GAGGATGTGGGATGCTGCTTGGGGCGGTGACCAACCACTGGGTCAGTTGGCC 1130
DB 724 GAGGATGTGGGATGCTGCTTGGGGCGGTGACCAACCACTGGGTCAGTTGGCC 783
QY 1131 TTTCGAGCCCTGAGACTCTTTCACAGCTTGAGATGGGAGACTGGAGCCCGGGAGAGGT 1190
DB 784 TTTCGAGCCCTGAGACTCTTTCACAGCTTGAGATGGGAGACTGGAGCCCGGGAGAGGT 843
QY 1191 GTGCTATTGTACCCCACTGCTCTACCGGCTGGGGGCTCTCTGCTAGAAAGGAGC 1250
DB 844 GTGCTATTGTACCCCACTGCTCTACCGGCTGGGGGCTCTCTGCTAGAAAGGAGC 903
QY 1251 TTTCACCACTGGGATGCTCCGGGGAGAGAGCTGA 1286
DB 904 TTTCACCACTGGGATGCTCCGGGGAGAGAGCTGA 939

```

RESULT 2
 US-08-176-427B-5
 ; Sequence 5, Application US/08176427B
 ; Patent No. 5789543

```

; GENERAL INFORMATION:
; APPLICANT: Ingham, Phillip W.
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Tablin, Clifford J.
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
; TITLE OF INVENTION: Proteins and Uses Related Thereto
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/176,427B
; FILING DATE: 30-DEC-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMI-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1056 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1056
; US-08-176-427B-5

Query Match 52.0%; Score 843.2; DB 2; Length 1056;
Best Local Similarity 88.2%; Pred. No. 5.5e-178;
Matches 928; Conservative 0; Mismatches 123; Indels 1; Gaps 1;

```

```

Db 361 GCCGCTGCAAGACAGCTGCTCTTCCCTGCCGAGCCCAAGTGGCCCTAGAGAAAGG 420
QY 696 GCGCGTGTGCGCTTGTACACCCGAGCCGAGACCGTGTGTGGCCATGGGGAGAT 755
Db 421 GAGCGTGTGCGCTGTACAGCTGTAAAGCAGAGACCGGGTGTGCGCATGGGGAGAT 480
QY 756 GGGAGCCCACTTACAGCATGTGTCTATTTTCTTGAGACCGGACCCCAAGAGCTGAGA 815
Db 481 GGGAGCCCACTTACAGCATGTGTCTATTTTCTTGAGACCGGACCCCAAGAGCTGAGA 540
QY 816 GCGTTCAGGATAGAGACTGAGACCCCGGCGGCTGAGCTACACCCGCTGAC 875
Db 541 GCTTTCAGGATAGAGACTGAGACCCCGGCGGCTGAGCTACACCCGCTGAC 600
QY 876 CTGCTCTTACAGGCTGACATACAGAGAGCCGAGCCGCTTCCGGGCAATTTGGC 935
Db 601 CTGCTCTTACAGGCTGACATACAGAGAGCCGAGCCGCTTCCGGGCAATTTGGC 660
QY 936 AGCCAGCTGACAGCTTGGCAGTACGTGTGTGGTGGGCTGCCAGGCTGACCTGCC 995
Db 661 AGCCAGCTGACAGCTTGGCAGTACGTGTGTGGTGGGCTGCCAGGCTGACCTGCC 720
QY 996 CGGCTGAGAGCTGTCTACACAGCTGGGCGCTGAGGCGCTAGCGCCCGCTCAAGAACAT 1055
Db 721 CGGCTGAGAGCTGTCTACACAGCTGGGCGCTGAGGCGCTAGCGCCCGCTCAAGAACAT 780
QY 1056 GGGACACAGTGTGTGAGATGTGTGGATCCCTGCTGCGGCGGCTGAGCAGCAC 1115
Db 781 GGGACACAGTGTGTGAGATGTGTGGATCCCTGCTGCTTGACACTGTGGTGTGACACAT 840
QY 1116 CTGCTGCTGCTGCTTGTGGCCCTGAGACTTTTCAAGCTTGGCAGTGGGAGTGGG 1175
Db 841 CTGCTGCTGCTGCTTGTGGCCCTGAGACTTTTCCAGTTGGCATGGGCAAGCTGG 900
QY 1176 ACCCGGGGAGGCTGTGATTTGATCCCGCACTGCTTACAGGCTGGGCGCTTCTCTG 1235
Db 901 ACCCGAGAGGAGGCTGTGATTTGATCCCGCACTGCTTACAGGCTGGGCGCTTCTCTG 960
QY 1236 CTGAGAGAGGAGGCTTCCACCCAGTGGGATGTCCGGGAGAGGAGTGAAGAGCTC- 1295
Db 961 CTGAGAGAGAGGAGCTTCCACCCAGTGGGATGTCTGGGGAGAGGAGTGAAGAGCTC 1020
QY 1295 CACCGCTGCCCTCTGGAAGTGTGCTGCTGAGG 1326
Db 1021 AACCACTGCCCTCTGGAAGTGTGCTGCTGAGG 1052

```

RESULT 3

```

US-08-356-060A-3
; Sequence 3, Application 5/08356060A
; Patent No. 5844079
; GENERAL INFORMATION:
; APPLICANT: Ingham, Phillip W.
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Tablin, Clifford J.
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356, 060A
; FILING DATE: 14-DEC-1994

```

```

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/176,427
; FILING DATE: 30-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMI-006CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1056 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1056
; US-08-356-060A-3

```

Query Match 52.0%; Score 843.2; DB 3; Length 1056;
Best Local Similarity 88.2%; Pred. No. 5,5e-178;
Matches 928; Conservative 0; Mismatches 123; Indels 1; Gaps 1;

```

QY 276 GAGCGCTTCAAGAGAGTCAACCCCAATTCATCCAGACATCATCTTCAAGAGAGAG 335
Db 1 GAGCGCTTCAAGAGAGTCAACCCCAATTCATCCAGACATCATCTTCAAGAGAGAG 60
QY 336 AACACAGGCGCGCAGCGCTCATGACCCAGCGCTGCAAGAGACCGCTGAAGTGGCTG 395
Db 61 AACACAGGCGCGCAGCGCTCATGACCCAGCGCTGCAAGAGACCGCTGAAGTGGCTG 120
QY 396 ATCTGCTGATGAACAGTGGCGCGGTGTGAAGCTGCGGCTGACCGAGCGCTGGAGAG 455
Db 121 ATCTGCTGATGAACAGTGGCGCGGTGTGAAGCTGCGGCTGACCGAGCGCGGTGA 180
QY 456 GAGCGCACACCTCAAGAGAGTCCCTGCAATATGAGAGGCGCGCGGTGAGACATCACCA 515
Db 181 GATGCTGATCACTCAAGAGAGTCTTACATATGAGAGGCGCGCGGTGATATCACAC 240
QY 516 TCAGACCGGAGCGCAATAGTATGAGAGTGGCGCGCTTGGCAGTGGAGCGCGCTT 575
Db 241 TCAGACCGGAGCGCAATAGTATGAGAGTGGCGCGCTTGGCAGTGGAGCGCGCTT 300
QY 576 GACTGGGTATATACAGATCAAGAGGCCCAAGTGCATTTGCTCCGTCAAGTCCGAGCACTG 635
Db 301 GACTGGGTATATACAGATCAAGAGGCCCAAGTGCATTTGCTCCGTCAAGTCCGAGCACTG 360
QY 636 GCGCGACCCCAAGAGCGCGGCTCTTCCCTCGGAGCCAGAGTACGCTGAGAGATGG 695
Db 361 GCGCGACCCCAAGAGCGCGGCTCTTCCCTCGGAGCCAGAGTACGCTGAGAGATGG 420
QY 696 GCGCGTGTGCGCTTACAGCGAGGAGCGGAGAGCGTGTGTGGCCATGGGGAGAT 755
Db 421 GAGCGTGTGCGCTTACAGCGAGGAGCGGAGAGCGGAGAGCGGAGAT 480
QY 756 GGGAGCCCACTTACAGCATGTGTCTATTTTCTTGAGACCGGACCCCAAGAGCTGAGA 815
Db 481 GGGAGCCCACTTACAGCATGTGTCTATTTTCTTGAGACCGGACCCCAAGAGCTGAGA 540
QY 816 GCGTTCAGGATAGAGACTGAGACCCCGGCGGCTGAGCTACACCCGCTGAC 875
Db 541 GCTTTCAGGATAGAGACTGAGACCCCGGCGGCTGAGCTACACCCGCTGAC 600
QY 876 CTGCTCTTACAGGCTGACATACAGAGAGCCGAGCCGCTTCCGGGCAATTTGGC 935
Db 601 CTGCTCTTACAGGCTGACATACAGAGAGCCGAGCCGCTTCCGGGCAATTTGGC 660
QY 936 AGCCAGCTGACAGCTTGGCAGTACGTGTGTGGTGGGCTGCCAGGCTGACCTGCC 995

```

Db	661	AGCCATGTGCACACCGGCCAATATATGTGTCTGTATCAAGGGGTACACAGGCGCTCCAGGCTCT	720
QY	996	CGCGTGGCAGCTGTCTCTACACACAGTGTGCGCTCTGGGGGCTACGCCCGCTCACAAAGCAT	10555
Db	721	CGGGTGGGAGGTGTCTCTCACACACAGTGTGCGCTCTGGGGTCTCTATGTCTCTCTTCACAAAGGAT	780
QY	1056	GGAGACCTGGGTGGGAGAGATGTGTGTGGCATCTCGCTTGGGGCGGTGGCTACACACAC	1115
Db	781	GGGACACTGTGTGTGGAGAGATGTGTGTGCTCTCTGTCTTGGACAGTGTGTGTGTACACACAT	840
QY	1116	CTGGCTCAGTTGTGCGCTTCTGGCGCTGTGAGACTTTTCAACAGCTTGGCATGTGGGACGTGG	1175
Db	841	CTGGCTCAGTTGTGCGCTTCTGGCGCTGTGAGACTTTTCCAGTTTGGCATGTGGGACGTGG	900
QY	1176	ACCCGGGGGAGGGGTGTGCATTTGGTACCCCACTGCTCTACCGCTGTGGGGGCTCTCTG	1235
Db	901	ACCCCAAGTGTGGGTGTGTACTCTCTACCTCTCAAGATGCTCTACCGCTGTGGGGGCTCTCTG	960
QY	1236	CTAGAAAGAGGACACTTCCACACCACTGTGGCATGTCCGGGGACAGAGAGCTGAAGAGACTC-	1295
Db	961	CTAGAAAGAGACACTTCCATCCACCTAGGGGATGTCTGTGGGGACAGAAAGCTGAAGGAGACTC	1020
QY	1295	CACCGCTGCCCTCTCTGGAATGCTGTACTGTGG	1326
Db	1021	AACCACTGCCCTCTCTGGAATGTGTGTGTGG	1052

```

1      RESULT      4
2      US-08-176-427B-1
3      ; Sequence 1, Application US/08176427B
4      ; Patent No. 5789343
5      ;
6      GENERAL INFORMATION:
7      APPLICANT: Ingham, Phillip W.
8      APPLICANT: McMahon, Andrew P.
9      APPLICANT: Tablin, Clifford J.
10     TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
11     TITLE OF INVENTION: Proteins and Uses Related Thereto
12     TITLE OF INVENTION: 33
13     NUMBER OF SEQUENCES: 3
14     CORRESPONDENCE ADDRESSES:
15     ADDRESSEE: LAHIVE & COCKFIELD
16     STREET: 60 State Street
17     CITY: Boston
18     STATE: MA
19     COUNTRY: USA
20     ZIP: 02109
21     ;
22     COMPUTER READABLE FORM:
23     MEDIUM TYPE: Floppy disk
24     COMPUTER: IBM PC compatible
25     OPERATING SYSTEM: PC-DOS/MS-DOS
26     SOFTWARE: ASCII(text)
27     ;
28     CURRENT APPLICATION DATA:
29     APPLICATION NUMBER: US/08/176,427B
30     FILING DATE: 30-DEC-1993
31     CLASSIFICATION: 435
32     ATTORNEY/AGENT INFORMATION:
33     NAME: Vincent, Matthew P.
34     REGISTRATION NUMBER: 36,709
35     REFERENCE/DOCKET NUMBER: HMT-006
36     TELECOMMUNICATION INFORMATION:
37     TELEPHONE: (617) 227-7400
38     TELEFAX: (617) 227-5941
39     INFORMATION FOR SEQ. ID NO.: 1:
40     SEQUENCE CHARACTERISTICS:
41     LENGTH: 1277 base pairs
42     TYPE: nucleic acid
43     STRANDEDNESS: both
44     TOPOLOGY: linear
45     MOLECULE TYPE: cDNA
46     FEATURE:
47     NAME/KEY: CDS
48     LOCATION: 1..1277
49     US-08-176-427B-1

```

Query Match	28.38;	Score 459;	DB 2;	Length 1277;
Best Local Similarity	65.28;	Pred. NO. 3.9e-93;		
Matches 763;	Conservative	0;	Mismatches 360;	Indels 48;
			Gaps	4;

QY	163	GGCACCGCCAGCCAAACCTGCTGCGCTGCGCTACAAAGAGTTGACGCCCATGCGCG	222
Db	107	GGAGGACCCCAAAAAGCTGACCCCGTTAGCCTATTAAGAGTTATTTCCAATGGGAG	166
QY	223	AGAAAGCCCTGGCGCCAGCGACGCTATGAAGGCAAGTCGCTCGCACGCTCGAGCGCT	282
Db	167	AGAAAGCCCTAGGGGGCCAGTGGAAAGTTAAGAGGGAAGTCCAAAGAACTCCGAGAT	226
QY	283	TCAAGAGCTCACCCCAATTACATCCAGACATATCTTCACAGGACGAGAGAAACAG	342
Db	227	TTAAAGAACTAACCCCAATTACACCCGACATTTATTTTAAAGATBAAAGAAACAG	286
QY	343	GCGCCGACCGCTCATATGACCCAGCGCTGCAGAGACCGCCTGAATCGCTGCTATCTGG	402
Db	287	GAGCTGACAACTGATGACTACAGCGCTGCAGAGACAGCTGATCCCTGGCATCTCG	346
QY	403	TGATGACCAAGTACCGCCGGGTGAACCTCGCGGTGACCGAGGCGTGGAGACGAGCCG	462
Db	347	TGATGACCAAGTACCGCCGGGTGAACCTCGCGGTGACCGAGGCGTGGAGAGATGGCC	406
QY	463	ACCACTCAGAGAGTCCCTGCAATTATGAGGCGCCGCGGTGACATCCACATCAGACC	522
Db	407	ATCATTCCGAGAGATCGCTCACACTACAGAGGTCGGCGGTGACATCCACAGTCGGATC	466
QY	523	GCGACCCCAATTAAGTTATGAGCTGCTGGCGCGTGTGACATGGAGGCGCGCTTGACTGG	582
Db	467	GGGACCGCACCAAGTACGGAATCTGCCCCGCTCGCGCTGAGAGCCGCGCTTCACATGG	526
QY	583	TGATTACAGAGTCAAAAGGCCACAGTCGATTCGCTCCGTCAGTCCGAGCACTGGCCGAC	642
Db	527	TCTACTACGAGTCCAAAGGCGCACATCACTGCTCCGTCAAAGCAAAACTCAGTGGCAG	586
QY	643	CCAAGACGGGCGGCTGCTTCCGTCGGCGGAGGCCAGTACCGCTGGAGATGGGGCGGTG	702
Db	587	CGAATTAGAGAGGTGCTTCTCCGCTCAGCCACAGATGACTCATTGGACATGAGGACCA	646
QY	703	TGCGCTTGTACCGCTGAGAGCGCGGAGACGTTGCTGTGCCATGGGGAGATGGAGCC	762
Db	647	AGCTGGTGAAGAGCACTGAGGCCCTGGGGACCGCGTGTGCTCTACCCGAGACGGCGG	706
QY	763	CCACCTTACGCGATGTGCTCATTTTCTGTGAGACCGGAGGCCACAGGCTGAGAGGCTTCC	822
Db	707	TGCTCTACATGTACTTCTCTACCTTCTCTGACCGGATGACAGCTCCCGAAGCTTCT	766
QY	823	AGGTCACTGAGACTTCAGAGACCCCCACAGCGCGCTGGCAATCACACCCGCTACCTGCTCT	882
Db	767	ACGTCACTGAGAGCGCGGACGCCCGGCGCGGTGTACTAGCAGCGGCCACACGCTCT	826
QY	883	TTAAGCGTCAATCAACAGGAGCCGG-----CAGCCCGCTTCGGGGCAAT	930
Db	827	TTTGGGCCCCCAGCACAAACCACTGTGGAGGCCAAGGGTCCACCAAGTGGACGAGGGCTT	886
QY	931	TTTGCACCACTGTCAGAGCTGGCCA-----GTAGCTGTGTGCTGGGGTGGCCAGGCC	984
Db	887	TCGCGACCAAGTGAAGCTGGCCCAACGTTGTTATGTGTGGGAGGGCGGGGACGAGC	946
QY	985	TGCAGCTGCGCGCGGTGGCAGCTGTCTTACAC---ACGTGGCCCTCGGGGCTTACGCC	1044
Db	947	TGCTGCGGGGTGTGTCACAGGCTCAATTCGGAGGAGGCGTCCGAGACTTACGCC	1006
QY	1042	CGCTCAAAAGCAATGGGACACTGGTGTGGAGATGTGTGGCACTCTGCTCGCGCGG	1102
Db	1007	CACCTACCGCCAGGAGCACTCTATATCAACCGGTTGTGGCTCTCGTACAGCGCTCA	1066
QY	1102	TGCGTACCAACCACTGGCTCACTGGCCTTGGCCCTCGAGACTCTTTACAGCTGG	1161
Db	1067	TCGAGGAGCAAGTTGGGCCATTTGGCCCTTGGCAACCAATTCGCTGGCTCAGGGGCTGC	1126


```

QY 390 CTGGTATCTCGGTATGATGAACAGTGGCCGGTGTGAACTGGGGTGAACCGAGGCTGG 449
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 476 TTGGCATCTCGGTATGAATACAGTGGCCAGAGGAACTCGGGTGAACCGAGGCTGG 535
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 450 GACGAGGAGCGCCACACACAGAGAGTCCCTGCAATAGAGAGGCGCGCGGTGAGCATC 509
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 536 GACGAGAGAGCGCCACACAGAGAGTCTGCACTAGAGAGGCGCGCGGTGAGCATC 595
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 510 ACCAGATCAGACCGGAGCGGCAATAGTATGAGTGTGGCGCTGGCTGGAGTGAAGGCC 569
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 596 ACCAGTGTGACCGGAGCGGAGCAAGTACGCGCATGCTGGCGCTGGCGGTGAGAGCC 655
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 570 GCGTTGACTGGGTATTATGAGTCAAAAGGCCACGTGCTGCTGCTGCTGCTGCTGCTG 629
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 656 GCGTTGACTGGGTATTATGAGTCAAAAGGCCACATATGCTGCTGCTGCTGCTGCTGCTG 715
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 630 CACTGGGCGGAGCAAGAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 689
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 716 AACTGGGTGGCGGCAATCGGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 775
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 690 AAGTGGGCGGCTGTGCTGTGCAAGCGGTGAGCGGAGCGGAGCGGCTGCTGCTGCTGCTG 749
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 776 CAGGCGGAGCAACACTGTGTGAAGAGCTGAGACCCCGGAGACCGGCTGCTGCTGCTGCTG 835
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 750 GAGATGGAGACCCCACTTCAGCGATGTCTCATTTTCTGAGACCGGAGCGGCAAG 809
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 836 GACGAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 895
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 810 CTGAGAGCTTCCAGTATGAGACTGAGAGACCCCGGAGCGGCTGCTGCTGCTGCTGCTG 869
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 896 AAGAAGGCTTCTAGCTGTGAGAGCGGAGCGGAGCGGAGCGGCTGCTGCTGCTGCTGCTG 955
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 870 GCTCACTGCTCTTTACGCGTGTGACATCAGAGAGCGGAGCGGCTGCTGCTGCTGCTGCTG 929
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 956 GCGCACTGCTTGTGTGGC---GCCGCAACAGACTGGCGCACCGGAGAGCGGAGCGGAGCG 1012
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 930 TTTCGCAACGACGTCAGCGCTGGCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 978
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 1013 TCCCTGGGCTGGGCGCGCTTCCGCGGAGCGACATGCGGCGCTTCCGCGGCGC 1061
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||

```

```

RESULT 9
US-08-176-427B-7
Sequence 7, Application US/08176427B
Patent No. 5789543
GENERAL INFORMATION:
APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tablin, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESSES:
ADDRESS: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/176.427B
FILING DATE: 30-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMI-006

```

```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1313 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1314
US-08-176-427B-7

```

```

Query Match 27.1%; Score 438.8; DB 2; Length 1313;
Best Local Similarity 65.1%; Pred. No. 1,1e-88;
Matches 704; Conservative 0; Mismatches 357; Indels 21; Gaps 3;
QY 91 TGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 150
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 32 TGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 91
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 151 TGGCAGCGCGCGGCGAGACCGGCAAACTGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 210
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 92 TTGGAAGAGCGGCGACCCCAAA---AAGTGACCCCTTACCTACAGAGATTATTC 148
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 211 CCAATGTCCCGAAGAAACCTTGGGCGGCGAGGAGCTGTGAAGGCAAGATCCCTGCA 270
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 149 CCAAGTATGCGGAGAAACCTTGGGCGGCGAGGAGCTGTGAAGGCAAGATCCCAAGAA 208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 271 GCTCGAGCGGCTTCAAGAGAGCTCAACCCCAATTCATCAAGACATCATCTCAAGAGC 330
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 209 ACTCGAAGCATTTAAGAACTACCCCAATTCATCAAGACATCATCTCAAGAGATG 268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 331 AGAAGAACAGCGCGGCGAGCGGCTCATGACCCAGCGCTGCAAGAGACCGGCTGAATGCG 390
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 269 AGAAGAACAGCGGCGAGCGGCTCATGACCCAGCGCTGCAAGAGACCGGCTGAATGCT 328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 391 TGGCTATCTGCTGTATGACCAAGTGGCGGCTGTGAAGTGTGGGTTGAGGAGGCTGGG 450
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 329 TGGCATCTCTGTATGATGACCAAGTGGCGGCTGTGAAGTGTGGGTTGAGGAGGCTGGG 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 451 ACGAGAGCGGCGACCTCAGAGAGTCCCTGCTATGAGAGCGGCGGCTGGAGATCA 510
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 389 ATGAGAGCGGCGACATTCAGAGAGTCTGACATGAGAGGCTGAGAGTGAAGTCA 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 511 CCACATCAGACCGGCGGCAATGATGAGTGTGAGGCGGCTGGAGTGGAGCGG 570
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 449 CCAGTCCGAGCGGCGGCGAGCAAGTACGCGTGTGCTGCTGCTGCTGCTGCTGAG 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 571 GCTTGAAGTGTATTAAGAGTCAAGGCGGCGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTG 630
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 509 GTTTCGACTGGGTCTATGATGATCAAACTGCACTGCTGCTGCTGCTGCTGCTGCTGCTG 568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 631 ACTGCGCGGAGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 690
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 569 ACTCGGCGGCGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 628
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 691 GTGGGCGGCTGTGCTGTGAGCGGCTGAGCGGCGGAGACCGGCTGCTGCTGCTGCTGCTG 750
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 629 AGGCGGCGGCGGCGGCGGCGGCTGAGAGCTTACGCGGAGACCGGCTGCTGCTGCTGCTG 688
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 751 AGATGGAGCGGCGGCGGCTTACGAGTGTGCTCATTTTCTGAGACCGGAGACCCCAAGCG 810
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 689 ACCAGGCGGCGGCGGCGGCTTACGAGTGTGCTCATTTTCTGAGACCGGAGAGAGCGGCGCA 748
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 811 TGAGAGCGGCTTCAAGTATGAGAGTCAAGAGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTG 870
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 749 AGAAGTCTTCTATGATGAGAGCTGAGAGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTG 808
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 871 CTCACCTGCTCT---TTACGGCTGACATCAACAGGAGCGGCGGCGGCTGCTGCTGCTGCTG 927
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||

```


Db	809	CGACACGCTCTCTCTGAGGCCGCCACACGACCTGGGGGCCACGCCCGGGCCAAAGCGCC	868
QY	928	CATTGGCAACCACTGACAGCTTGGCCAGTACGTGCTGGGTGGGTGCCAGGCTTGC	987
Db	869	TCATTGGCAACCCGCGTGGCCGCCGCCGACACGCGTGTACCTGTGGCTTAAACCGCGCGGG	928
QY	988	AGCTGGCCCGCGTGGACGCTGTCTACACAGCTGGCCCT-----CGGGG	1032
Db	929	ACCGCCGCGTGTGCCCCGCCGCGGTACACAGCTGTACGCTGGCAGAGAGAGAGCGCGGCG	988
QY	1033	CGTACGCCCCCGCTCACAAAGCATGGGACACTGGTGGTGGAGGATGTGGGCATCTGCT	1092
Db	989	CGTACGCGCGCGCTACAGGGCGCACGGCACCATTTCTATCAACGGGGTGTGCTCGCTGTGCT	1048
QY	1093	TCGCGGCGGTGGCTGACCAACCACTTGCTGTAAGTTGGCCTTCTGGCCCTGAGACTTTTC	1152
Db	1049	ACGCTGTCACTGAGAGACACAGCTGGGACACACCGGCGCTTCGCGCCTTTCGCGTGGCGC	1108
QY	1153	AC	1154
Db	1109	AC	1110

```

RESULT 10
US-08-356-060A-4
Sequence 4, Application US/08356060A
Patent No. 5844079
GENERAL INFORMATION:
APPLICANT: Ingham, Phillip W.
APPLICANT: Mahon, Andrew P.
APPLICANT: Rablin, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
TITLE OF INVENTION: Proteins and Uses Related Thereto
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESSES:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,060A
FILING DATE: 14-DEC-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/176,427
FILING DATE: 30-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMT-006CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1313 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1314
US-08-356-060A-4

```

[illegible]

Db	280	GGGAGACGGCTCAGACACAGAGATGAAGAAAGAACACTGAACCTCGGTGGCATTCTCTTA	339
QY	405	ATGACCCAGTGGCCCCGGTGTGAAGCTGGGGTACCAGAGGGCTGGGACGAGCGGCAC	464
Db	340	ATGAACCACTGGCCACGGGGTTAAGACTCGGTGACAGAGAGGGCTGGGATGAGACGGTCAC	399
QY	465	CACTCAGAGGAGTCCCTGTCATTATGAGGGCGCGCGGTGACATCACCAATCAACAGCC	524
Db	400	CATTTTGAAGATATCTCCACTACGAGGGAAGAGCTGTTGATATTAACCACTCTACGA	459
QY	525	GACCGCATAAGTATGGACAGTGGCCGGCTTGAGAGTGAAGGCGGGCTTACTGGGTG	584
Db	460	GACAAAGACAAATACGGGAACTGTCTCCGCTTAGCTGTGAGAGCTTGGATTTACTGGGTC	519
QY	585	TATTACGAGTCMAAGGCCCACTGTGCATTGCTCCGTCAAGTCCGAGCACTCGGCCGACCC	644
Db	520	TATTACGAGTCMAAGGCCCACTTANTGTGTGTCAAAAGCAAGAAATTCGGTCTCGG	579
QY	645	AAGACGGGCGGCTCTTCCTGCCGGAGCCCAAGTACGCGCTGGAAGATGGGCGCGTGTG	704
Db	580	AAATCTGGGGCTCTTTTCCACAGGTTGGGCTCTGTCTCGCTCCAGGACGGAAGACAAG	639
QY	705	GCGTTGTACGCCGAGAGCCGGGAGAACCGTGTCTGGCCATGGGGAGATGGAGAGCCC	764
Db	640	GCGGTGAAGGACCTGAACCCCGGAGACAAAGTGTCTGGCGACAGACGCCGGGAAACTGG	699
QY	765	ACCTTCACGATGTGCTATTTTCTGTGAGCCGGAGCCCAAGGCTGTGAGACCTTTCAG	824
Db	700	GTTGTCAGGCACTTCATCATGTTCACAGACGAGACTCCACAGACGGAGGTGTGTTTAC	759
QY	825	GTCATCGAGACTCAGGAGCCCCCAAGCCGCGCTGGCACTACACCCGCTCACTGCTTT	884
Db	760	GTCATAGAAACGCAAGAACCCGTTGAAGAAATCACCTCATCCGCGCTCACTCCTTTT	819
QY	885	ACGGCTGACATCACACGAGACCGGACGCCGCTTCCGGGCCAATTTGCCAGCCACGTG	944
Db	820	GTCCTCGCAACTCAACGGAAGATCTCCACACCATACCGCGCGGTATCCACGAGTGTG	879
QY	945	CAGCGTGGCCAGTACGTGTGAGTGGGCTGGGGGTGCCAGGGCTCG---AGCCTGCCGGG	1001
Db	880	AGAGCCCGACAAAGGTGATGATGTTGTGATGATAGCGGTACAGTTAAATCTGTCAATCGTG	939
QY	1002	GCACCTGTCTTACACACGTGGGCTCTGGGGCTACAGCCCCGCTCAAAAGATGGAGACA	1061
Db	940	CAGGGAATATCACAGGAGAGACAGCGGGGCTGCTTGGCACAAGTACTCAATGGAGAC	999
QY	1062	CTGGTGTGAGAGATGTGTGGCATCTCTCGCGGCGGTGGCTATCCACCACTGGCT	1121
Db	1000	ATTGTGTGTCGACAAATCTGCGGCTCCTGTTCAGCGGTAAAGAGACCAAGGGGCTTGGC	1059
QY	1122	CAGTTGGCTTCTGGCCCCCGAGACACTTTCACAGCTTGGCAT	1164
Db	1060	CATTGGCTTCTGGCCCCCGAGGCTTATTATTACGTGTCAAT	1102

Search completed: June 5, 2000, 07:56:48
Job time: 3779 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2000, 06:55:15 ; Search time 74.45 Seconds
(without alignments)
5450.793 Million cell updates/sec

Title: US-08-900-220-7

Perfect score: 1622
Sequence: 1 CATCAGCCGACACGAGAGACC.....CATTCGAGGAGCCCATTCCTCC 1622

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1622	100.0	1622	1	X16188 Human Ihh hedgehog
2	1622	100.0	1622	1	X07269 Human Indian hedgehog
3	1622	100.0	1622	1	X25104 Human Indian hedgehog
4	1029.8	63.5	1281	1	X16184 Mouse Ihh hedgehog
5	1029.8	63.5	1281	1	X07273 Mouse Indian hedgehog
6	1029.8	63.5	1281	1	X25100 Mouse Indian hedgehog
7	936	57.7	939	1	O91641 Human Indian hedgehog
8	843.2	52.0	1056	1	O91640 Mouse Indian hedgehog
9	459	28.3	1277	1	O91636 Chicken sonic hedgehog
10	459	28.3	1277	1	X16182 Chicken Shh hedgehog
11	459	28.3	1277	1	X07271 Chicken sonic hedgehog
12	459	28.3	1277	1	X25098 Chicken sonic hedgehog
13	453.8	28.0	1190	1	X07270 Human Desert hedgehog
14	453.8	28.0	1190	1	X25097 Human Desert hedgehog
15	453.8	28.0	1425	1	O91639 Human Sonic hedgehog
16	453.8	28.0	1425	1	O91637 Human Shh hedgehog
17	453.8	28.0	1425	1	X07276 Human Sonic hedgehog
18	453.8	28.0	1425	1	X25103 Human Sonic hedgehog
19	452.2	27.9	1576	1	V18403 Human mutated sonic hedgehog
20	452.2	27.9	1576	1	V18404 Human mutated sonic hedgehog
21	450.4	27.8	1188	1	V62395 Human Desert hedgehog
22	442.4	27.3	1122	1	V62394 Human Desert hedgehog
23	438.8	27.1	1313	1	O91637 Human sonic hedgehog
24	438.8	27.1	1313	1	X16185 Human sonic hedgehog
25	438.8	27.1	1313	1	X07274 Mouse sonic hedgehog
26	438.8	27.1	1313	1	X25101 Mouse sonic hedgehog
27	430.8	26.6	1715	1	T04368 Mouse Sonic hedgehog
28	419.8	25.9	1190	1	O91642 Rat Vhh-1 cDNA, Nu
29	419.8	25.9	1190	1	X16182 Mouse desert hedgehog
30	419.8	25.9	1190	1	X07272 Mouse Dh hedgehog
31	419.8	25.9	1191	1	X25099 Mouse Desert hedgehog
32	418.2	25.8	1256	1	O91638 Zebrafish sonic hedgehog
33	418.2	25.8	1256	1	X16186 Zebrafish Shh hedgehog
34	418.2	25.8	1256	1	X07275 Zebrafish Sonic hedgehog

35	418.2	25.8	1256	1	X25102 Zebrafish Sonic hedgehog
36	369	22.7	1251	1	X16189 Zebrafish Tth hedgehog
37	369	22.7	1251	1	X07277 Zebrafish tlgie-w
38	369	22.7	1251	1	X25106 Zebrafish tlgie-w
39	341	21.0	522	1	V62400 Human Sonic hedgehog
40	318.4	19.6	530	1	V59456 Nucleotide sequence
41	307.6	19.0	602	1	V62397 Human Desert hedgehog
42	303.2	18.3	548	1	V62396 Human Desert hedgehog
43	297.6	18.3	528	1	V62393 Human Desert hedgehog
44	276.2	17.0	1416	1	X16190 Drosophila HH hedgehog
45	276.2	17.0	1416	1	X25107 Drosophila hedgehog

ALIGNMENTS

RESULT	ID	Sequence	Score	DB ID	Description
1	X16188	standard; cDNA; 1622 BP.	100.0%	DB 1	Length 1622;
	X16188		Best Local Similarity		
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
		Query Match	100.0%	Score 1622;	DB 1; Length 1622;
		Best Local Similarity	100.0%	Pred. No. 0;	
		Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			

ID	Accession	Gene	Location/Qualifiers
ID	X07273	standard; cDNA; 1281 bp.	
AC	X07273:		
DT	21-MAY-1999	(first entry)	
DE	Mouse indian hedgehog (Ihh) cDNA.		
KW	Indian hedgehog; Ihh gene; mouse; dopaminergic; Parkinson's disease; ptc therapeutic; patched; signal transduction; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; cerebral ischemia; hypoxia; neuroprotective; gene therapy; ss; ds.		
OS	Mus sp.		
FH	Key	Location/Qualifiers	
FT	CDs	1..1236	
FT		/*tag= a	
PD	04-FEB-1999.		
PF	24-JUL-1998; U15419.		
PR	24-JUL-1997; US-900220.		
PA	(ONTO-) ONTOGENY INC.		
PI	Mahantappa NK, Miao N, Pang K, Wang M;		
DR	WPI:99-142578/12.		
DR	P-PsDB: W97767.		
PT	Increasing the survival of neuronal, dopaminergic and GABA-nergic cells - by using a ptc therapeutic such as a protein kinase inhibitor, or an agent derived from hedgehog polypeptides, useful in the treatment of Parkinson's disease		
PT	Disclosure: Page 87-89; 138pp; English.		
PS	This nucleotide sequence includes a coding region for the mouse Ihh indian hedgehog protein (see W97767). The invention is based on the finding that hedgehog proteins are useful as protective agents in the treatment and prophylaxis of neurodegenerative disorders resulting from the loss of dopaminergic and/or GABA-nergic neurons, or the general loss of tissue from the substantia nigra.		
CC	Exemplary disorders include Parkinson's disease, Huntington's disease (both claimed), amyotrophic lateral sclerosis and cerebral ischemia. The invention relates to hedgehog therapeutics (i.e. hedgehog polypeptides and gene therapy constructs e.g. constructs encoding recombinant hedgehog polypeptides and trans-activation constructs for altering hedgehog gene regulatory sequences) and ptc therapeutics (i.e. agents which mimic the effect of naturally occurring hedgehog proteins on patched signalling) that are effective in both human and animal subjects. Human Ihh and Dhh polypeptides (see W97763-64) are preferred. The products can also be used for the maintenance of differentiated neurons in cultures, and to enhance the implantation of such neuronal cells in an animal. They can also be used to prevent or treat neurodegenerative conditions arising from the use of certain drugs, and in the prevention and/or treatment of hypoxia, e.g. as a neuroprotective agent.		
CC	Sequence 1281 bp; 229 A; 401 C; 393 G; 258 T;		
SO			
Query Match	63.5%; Score 1029.8; DB 1; Length 1281;		
Best Local Similarity	88.4%; Pred. No. 2.9e-194;		
Matches 1129; Conservative	0; Mismatches 147; Indels 1; Gaps		
QY	51 ATGTCTCCGCGCGGCTCCGGCCCGAGTCGACTTGTGCTGTGCTGTGCTGTG 110		
DB	1 ATGTCTCCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 60		
QY	111 GGTGTGCGCGCGCGCATGGGGCTGTGGGGCGGGTCTGGGTGTGGGACGCGCGCGACCG 170		
DB	61 CTGTGTGCGCGCGCGCGGGGCTGTGGGGCGGGCGGGTGTGGGCGAGCGCGGAGCGCG 120		
QY	171 CCAGCGAACTGTGCGCGGTGTGCAAGACAGTTAGCCCAAGTGGCCGAGAAAGACC 230		
DB	121 CTTGCGAACTGTGTGCTTGTGCTTCAAGACAGTTAGCCCAAGCTGCGGAGAAAGACC 180		
QY	231 CTGGGCGCGAGCGAGCTATGAGGCAAGATCGCTCGAGCTCCGAGCGCTTCAAGAGAG 290		
DB	181 CTGGGCGCGAGCGGGGCGCTTACGAGGCAAGATCGCGCGAGCTCTGAGCGCTTCAAGAG 240		
QY	291 CTCACCCCAATTACATCAATCAATCATCTTCAAGAGACAGAGAAACAGCGCGCGAC 350		

Db	241	CTCACCCCCCACTGACATGCCGACATCATCTTCACAGAGACAGGAGAAACAACGGGTGCCGAC	300
QY			
351	CGCCTCATGACCCACAGCGCTGCAGAGGACCGCCTGAACTCGCTGGCTATCTCGGTATGTAAC	410	
Db	301	CGCCTCATGACCCACAGCGCTGCAGAGGACCGCTGAACTCACTGGCATCTCTGTATGTAAC	360
QY			
411	CAGTGGCCCCGGTGTGAAGCTGGGGGTGACCCGAGAGGTGTGGGACGAGAGACGGCCACACCTCA	470	
Db			
361	CAGTGGCCCCGGTGTGAAGCTGGGGGTGACCCGAGAGGTGTGGGACGAGAGAGTGGCCATCACTCA	420	
QY			
471	GAGAGATGCCCGCATTTATGAGGGCCCCCGGGTGGACATATCACATCATGACCGGACCGC	530	
Db	421	GAGAGATCTTACACTATGAGGGCCCCCGGGTGGATATACCACTCATGACCGTGTACCGA	480
QY			
531	AATAAGTATGAGACTGCTGGCGGCCCTTGGACATGAGAGCGCGCTTTAGCTGGGTATTTAC	590	
Db	481	AATAAGTATGAGACTGCTGGCGGCCCTTAGCAGTGGAGAGCGCGCTTGCATGGGTGATTTAC	540
QY			
591	GAGTCGAAGGCCACACGTGCAATTGCTCCGTCAAGTCCGAGCACTCGGGCCGACCCAAAGC	650	
Db			
541	GAGTCGAAGGCCACACGTGCAATTGCTCTGTCAAGTGTGACCAATTCCGGCCGCTGCCAAAGCA	600	
QY			
651	GCGCGCTGCTCCCTCGCCGGAGCCCAAGATACGCGCTGAGAGTGGGGGCCGCTGTGGACCTTG	710	
Db	601	GCTGGCTGCTTTCTCTGCCGAGACCCAGGTGCGCTTAGAGAACGGGAGACGTGTGGCCCTG	660
QY			
711	TCAACCGGTGAGGCCGGGAGACCCGTGTGCTGTGGCCATGTGGGGAGATGGGAGGCCCACTTC	770	
Db	661	TCACCTGTAAAGCCAGGAGACCGGGTGTGGCCATGTGGGGAGATGGGAGGCCCACTTC	720
QY			
771	AGCGATGTGCTCATTTTCTGTGACCGCCGAGCCCCACAGGCTGAGAGCCTTCCAGGTATC	830	
Db			
721	AGTATGTGCTTATTTTCTGTGACCCCGAGCCAAACCGGCTGAGAGCTTCCAGGTATC	780	
QY			
831	GAGACTGAGGACCCCCCAGCGCGCGCTGGGACACTACACCCGCTACCTGCTTTTAGGCT	890	
Db	781	GAGACTGAGGATCTCCGCGCTCGGCTGGGCTCAGCTCAGCCCTGCCACCTCTCTTCACTTGG	840
QY			
891	GACATCATACACGAGCCGGGACGCGCCCTTCCGGGGCACATTTGGCCAGCACAGTGGACGCT	950	
Db	841	GACATCATATACAGAACGAGGACCCCACTTCCGGGCCACATTTGCCAGGCATGTGGAAACA	900
QY			
951	GGCCAGTACGCTGCTGTGGCTGGGGGTGCCAGGCTGTGACGCTGCCCCGCTGGCAGCTGTC	1010	
Db	901	GGCCAAATATGTGCTGTATCAGGGGTATCCAGAGGCTCTCAGACCTGTGCGGGTGGCAGCTGTC	960
QY			
1011	TCTACACAGCGGGCCCGGGGGGCTACGCGCCGCTACAAACACATGGGACACTGTGTTG	1070	
Db	961	TCACCCACCGTGGCCCTTGGGTCTATAGCTCTCTCTACAAAGGCAATGGGACACTGTGTTG	1020
QY			
1071	GAGAGTGTGATGAGCATCTGCTTCGAGGCGCGTGGGCTGACACACACTGGCTCAGTTGGCC	1130	
Db	1021	GAGATGTGGGTGGGCTCTGCTTTGCAAGCTGTGGGTGTGACACACACTGGCTCAGTTGGCC	1080
QY			
1131	TTCTGGCCCCCTGAGACTCTTTTCACAGCTTGGCATGGGGCAGCTGGACCCCGGGGAGGGT	1190	
Db	1081	TTCTGGCCCCCTGAGACTCTTTCCCACTTTGGCATGGGGCAGCTGGACCCCAAGTAGGGT	1140
QY			
1191	GTGCAATGTGTAACCCCAAGCTGCTTACCGCTGTGGGGCTCTCTCTGTAGAGAGGGCAGC	1250	
Db	1141	GTTCATCTCTACCTCTGAGATGCTCTACCGCTGTGGGGCTCTCTCTGTAGAGAGGACCC	1200
QY			
1251	TTTCACCCACTGGGACATGTCCGGGGGAGGAGAGTGAAGAGATC-CACCGCTGCCCTGCT	1309	
Db	1201	TTTCATCTCACTGGGACATGTGTGGGGGAGAGAGCTGAAGGAGACTTAACCACTGCCCTCTCT	1260
QY			
1310	GGAAGTCTGTACTGGG 1326		
Db	1261	GGAAGTCTGTGTGGTGG 1277	

ID	X25100: standard; cDNA, 1281 BP.
DT	X25100:
DE	Mouse Indian hedgehog protein 1bh cDNA.
KW	Indian hedgehog; 1bh gene; mouse; hedgehog therapeutic;
KW	ptc therapeutic; patched; signal transduction; muscle atrophy;
KW	cachexia; muscular myopathy; myoblastic sarcoma; therapy; ds.
OS	Mus sp.
FT	
FT	Key
FT	1..1236
FT	/*tag= a
PN	W09910004-A2.
PD	04-MAR-1999.
PF	28-AUG-1998: U17922.
PR	29-AUG-1997: U5-057394.
PA	(ONTO-) ONTOGENET INC.
PI	Bladen CS, Currie PD, Hughes SM, Ingham PW;
DR	WPI: 99-243557/20.
DR	P-PSDB: Y05512.
PT	A new method to regulate muscle growth
PS	Disclosure: Page 96-98; 130pp; English.
CC	This nucleotide sequence comprises a coding region for the mouse
CC	Indian hedgehog protein 1bh (see Y05512). The invention relates to
CC	a method for modulating the formation and/or maintenance of muscle
CC	tissue by ectopically contacting muscle cells, especially muscle
CC	stem/progenitor cells, in vitro or in vivo, with a hedgehog
CC	therapeutic (i.e. hedgehog polypeptides and gene therapy
CC	constructs) or ptc therapeutic (i.e. a small organic molecule that
CC	mimics the effect of hedgehog proteins on patched signalling, or
CC	activates or potentiates patched signalling) in an amount effective
CC	to alter the growth state of the treated cells. Also claimed is a
CC	method for treatment or prevention of disorders of, or surgical or
CC	cosmetic repair of, such muscle tissues, by administering a
CC	hedgehog polypeptide or ptc therapeutic. The disorder may be
CC	muscle atrophy, in particular skeletal muscle atrophy or cardiac
CC	muscle atrophy, cachexia, or muscular myopathy (all claimed). The
CC	hedgehog polypeptide or ptc therapeutic can inhibit growth of
CC	myoblastic-derived tissue to provide treatment of hyperblastic or
CC	neoplastic growth of muscle tissue such as in myoblastic sarcoma
CC	(also claimed). The hedgehog therapeutic preferably comprises at
CC	least a bioactive extracellular portion of a hedgehog protein (see
CC	Y05510-19) encoded by a vertebrate hedgehog gene (see X25098-107),
CC	especially a human hedgehog gene.
SO	Sequence 1281 BP: 229 A; 401 C; 393 G; 258 T;

Query Match	63.5%;	Score 1029.8;	DB 1;	Length 1281;
Best Local Similarity	88.4%;	Pred. No. 2.9e-194;		
Matches 1129; Conservative	0;	Mismatches 147;	Indels 1;	Gaps 1

QY	51	ATGATCCTCCGGCCGGGCTCCGGGCCCCGAGTGCATCTTTCGCGATGCTGTTGCTGCTG	110
Db	1	ATGATCCTCCGGCTCGGCTCCGGCCCCGAGTGCAGTCTGTTGCTGTTGCTGCTGCTG	60
QY	111	GTGGTCCCCGGGCATGGGGGCTGGGGGCTGGGGTGGTGGGGACCCCGGGCACCG	170
Db	61	CTGGTGGCGGGCGGGCGGGGCTGGGGGGCGGGCGGGGTGGGGACCCCGGGAGGCC	120
QY	171	CCACGCAAACTGTATCCGCTCGCTCCCTACAGAGAGTTTACAGCCCAATGTGCCGAGAAACC	230
Db	121	CCTGCGAAGCTGTATCCCTCTTGCTGCTACAGAGAGTTTACAGCCCAAGTGGCGAGAAACC	180
QY	231	CTGGGGCGCCAGCGGACGCTATGAAGGCAAGATCGCTGGCAGCTTCGAGCGCTTCAAGAG	290
Db	181	CTGGGGCGCCAGGGGGCGCTTACGAAGGCAAGATCGCGGACCTTGGAGCCCTTCAAGAG	240
QY	291	CTCACCCCCAATTACAAATCCAGACATCATCTTTCAAAGACGAGGAAACACAGCGCCGAC	350
Db	241	CTCACCCCCCAATTACAAATCCGAGATCATCTTTCAAAGACGAGGAAACAGCGGTCCGAC	300
QY	351	CGGCTCATGACCCAGCGCTGCAGAGGACCGCTGAACCTGCTGGCTATCTGGGTATATAC	410
Db	301	CGGCTCATGACCCAGCGCTGCAGAGGACCGGTGGAACCTCATCGGCAATCTGTGATATAC	360

OY	411	CAGTGGCCCGGTGAAAGCTGCGCGGTGACCGAAGGGCTGGGAGAGAGACCGCACACTCA	470
Db	361	CAGTGGCCCGGTGTGAAACACTGCGGGGTGACCGAAGGCCCGGGATGAAGATGGCCATTACTCA	420
OY	471	GAGGAGTCCCTGCAATATATGAGGGCCGCGGGGTGAGCATCACCATCAGACCGCACCGC	530
Db	421	GAGGAGCTTTTACACTATATGAGGGCCCGCGGGTGGATATCACCACTTCAGACCGTACCGA	480
OY	531	AATAAGTATGAGACTGCTGCGGGCGCCTTGGCATGAGAGGCCGCTTGGACTGGGTGATTAC	590
Db	481	AATAAGTATGAGACTGCTGCGGGCGCCTTGGCATGAGAGGCCGCTTGGACTGGGTGATTAC	540
OY	591	GAGTCAAGAGGCCCGACGTCATTTGCTCTCGTCAAGTCCGAGACATCTGGCGGACGCAAGAG	650
Db	541	GAGTCAAGAGGCCCGACGTCATTTGCTCTGTCGTAAGTCTGAGACATCTGGCGGCTGCCAAGCA	600
OY	651	GCGCGCTGCTTCCCTGCGGAGGCCACAGTATGCGCTGAGAGTGGGGCGGTGTGGCTTG	710
Db	601	GGTGGCTGCTTCCCTGCGGAGGCCACAGTATGCGCTTGGATGAGAAAGGGGAGGTGTGGCCTG	660
OY	711	TCACACCGTGAAGCCCGGGAAGACCGTGTGTGTGGCCATGTGGGGGAAGATGGAGCCCACTTC	770
Db	661	TCACACTGTAAAGCCGAGGAGCCGCGGTGTGGCCATGTGGGGGAAGATGGAGCCCACTTC	720
OY	771	AGCGATGTGCTATTTTCTTCCGAGACCCGACGCCACAGGCTGAGAGCCTCCAGGCTATC	830
Db	721	AGTGTATGCTTATTTTCTTCCGAGACCCGAGACCAACCGGCTGAGAGCTTTCAGAGTATC	780
OY	831	GAGACTGAGGACCCCGACCGCCGCTGGCACTACACCCGCTCACCTGCTTTTACGGCT	890
Db	781	GAGACTGAGGATCCCTCCGCGTGGGTGGGCTACGCGTCCCACTGCTTTTACGCT	840
OY	891	GACATCATCACGGAAGCCGGGACGCCGCTTCCGGGCCACATTGGCAAGCCACGTGCAAGCT	950
Db	841	GACATCATACAGAACGACGACGCCACTTCCGGGCCACATTGGCAAGCCATGTGCAACCA	900
OY	951	GGCCAGTACGCGTGGGTGGGGTGGGAGGCCAGGCGTCAACCTTGGCCCGGGTGGCACTGTC	1010
Db	901	GGCCAAATATGTGCTGGTATATAGGGGGTATACAGGCTTCAACCTTGTGGGTGGCACTGTC	960
OY	1011	TCTACACAGCTGGGCCCTCGGGGGCTCTAGGCCGCCGCTCACAAAGCATGGGACACTGGTGTG	1070
Db	961	TCACACCAAGCGGCCCTTGGGTCTATATGCTCTCTACAAAGCATGGGACACTGGTGTG	1020
OY	1071	GAGGATGTGGTGGCATCTCTCTTCCGGGCCGTGTGCTGACACCACTGGCTCAGTTGGCC	1130
Db	1021	GAGGATGTGGTGGCTCTCTCTTGGAGCTGTGTGGCTGACACCACTGGCTCAGTTGGCC	1080
OY	1131	TTTGGGCCCCGAGACCTCTTTCACAGCTTGGGATGGGAGAGTGGAGCCCGGGGAGAGGT	1190
Db	1081	TTTGGGCCCCGAGACCTGTTTCCAGTTTGGCATGGGGAGAGTGGAGCCCAAGTGAAGGT	1140
OY	1191	GTGCAATGTGTACCCCAAGCTCTCTACAGCGCTTGGGCGCTCTCTGCTGTAAAGAGGAGC	1250
Db	1141	GTTCACATCTACCCCTCAGATGTCTACCGCGCTTGGGCGCTCTCTGCTGTAAAGAGGAGC	1200
OY	1251	TTTCAACCACTGGGAGCTGTCCGGGGGACGGGAGCTGAAGGACTC-CACCGCTGCCCTCT	1309
Db	1201	TTTCAATCCACTGGGAGCTGTCCGGGGGACGGGAGCTGAAGGACTTAAACCACTGCCCTCT	1260
OY	1310	GGAACTGCTGTACTGG 1326	
Db	1261	GGAACTGCTGTGCTGG 1277	

RESULT	7
ID	O91641
AC	O91641 standard; cDNA; 939 BP.
DT	18-MAR-1996 (first entry)
DE	Human indian hedgehog protein gene.
KW	Human; Indian hedgehog gene; nested polymerase chain reaction; PCR;

KW fetal lung; probe; primer; diagnostic; nervous system disorder;
 KM gene therapy; antibody; ss.
 OS Homo sapiens.
 PN W09518856-A1.
 PD 13-JUL-1995.
 PF 30-DEC-1994; U14992.
 PR 30-DEC-1993; US-176427.
 PR 14-DEC-1994; US-356060.
 PA (HARD) HARVARD COLLEGE.
 PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
 PI Ingham PW, McMahon AP, Tabin CJ;
 DR WPI: 95-255060/33.
 DR P-PSDB: R77344.
 PT Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful
 PT to treat degenerative nervous system disorder(s) and in gene
 PT therapy.
 PS Claim 4; Page 146-47; 210pp; English.
 CC The sequence encodes a human Indian hedgehog protein, homologous
 CC to a Drosophila hedgehog protein (R77337), and has been isolated by
 CC screening of human genome DNA by nested polymerase chain reaction
 CC using primers 091643, 091644 and 091645, followed by use of a clone
 CC lambda gt-10. Probes and primers derived from hedgehog sequences
 CC may be used as diagnostic agents for neuromuscular, autonomic or
 CC central nervous system disorders, and the gene may also be used in
 CC gene therapy. Antibodies generated from the encoded protein may be
 CC used as therapeutic or research reagents.
 SQ Sequence 939 BP; 154 A; 305 C; 314 G; 166 T;

Query Match 57.7%; Score 936; DB 1; Length 939;
 Best Local Similarity 100.0%; Pred. No. 7.6e-176;
 Matches 936; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 351 GCGCTCATGACCCAGCGCTGCAAGAGCCGCTGAACCTGCTGGCTATCTCGGTATGAC 410
 DB 4 CGCCCATGACCCAGCGCTGCAAGAGCCGCTGAACCTGCTGGCTATCTCGGTATGAC 63
 QY 411 CAGTGGCCGGGTGAGTACCTGGGGTGACCGAGGCGTGGAGAGGAGGCGGACCACTCA 470
 DB 64 CAGTGGCCGGGTGAGTACCTGGGGTGACCGAGGCGTGGAGAGGAGGCGGACCACTCA 123
 QY 471 GAGGAGTCCCTGCTATATAGAGGCGCGCGGTGAGCATCACACATCAGACCGGACCGC 530
 DB 124 GAGGAGTCCCTGCTATATAGAGGCGCGCGGTGAGCATCACACATCAGACCGGACCGC 183
 QY 531 AATAAGTATGACTGCTGGCGGCTTGGCACTGAGAGCGCGGCTTGAAGTGGTATTAC 590
 DB 184 AATAAGTATGACTGCTGGCGGCTTGGCACTGAGAGCGCGGCTTGAAGTGGTATTAC 243
 QY 591 GAGTCAAGGCGCCAGCTGATTCGCTCAAGTCCGAGCATCGCGGCGGACGCAAGCG 650
 DB 244 GAGTCAAGGCGCCAGCTGATTCGCTCAAGTCCGAGCATCGCGGCGGACGCAAGCG 303
 QY 651 GCGGCGCTCTTCCCTGCGGAGCCAGGTAGCGCTGAGAGTGGGGCCGTGTGGCTTG 710
 DB 304 GCGGCGCTCTTCCCTGCGGAGCCAGGTAGCGCTGAGAGTGGGGCCGTGTGGCTTG 363
 QY 711 TCAGCGGTGAGGCGGAGAGCGGTGCTGGCATAGGGGAGAGTGGAGCCCACTTC 770
 DB 364 TCAGCGGTGAGGCGGAGAGCGGTGCTGGCATAGGGGAGAGTGGAGCCCACTTC 423
 QY 771 AGCGATGCTCATTTTCTTGAGCGAGGCGCCAGCGAGAGAGGCTTCCAGGTATC 830
 DB 424 AGCGATGCTCATTTTCTTGAGCGAGGCGCCAGCGAGAGAGGCTTCCAGGTATC 483
 QY 831 GAGACTCAGAGCCGCCAGCGGCTGGCATCAGACCCGCTCACTCTCTTTACGGCT 890
 DB 484 GAGACTCAGAGCCGCCAGCGGCTGGCATCAGACCCGCTCACTCTCTTTACGGCT 543
 QY 891 GACATATCAGAGCGGAGCGGAGCGGCTTCCGGGCGACATTTGCCAGCGAGTGCAGCT 950
 DB 544 GACATATCAGAGCGGAGCGGAGCGGCTTCCGGGCGACATTTGCCAGCGAGTGCAGCT 603

QY 951 GCGCAGTACGTCGTCGTGGGTGGGTGCGCAGGCGCTGAGCGCCGCGGAGCTGTC 1010
 DB 604 GCGCAGTACGTCGTCGTGGGTGGGTGCGCAGGCGCTGAGCGCTGCGCGGAGCTGTC 663
 QY 1011 TCTACACAGCTGGGCGCTGAGGCGCTTACGCGCCGCTCAAAAGCATGGAGACGTGTG 1070
 DB 664 TCTACACAGCTGGGCGCTGAGGCGCTTACGCGCCGCTCAAAAGCATGGAGACGTGTG 723
 QY 1071 GAGATATGCTGTCGTCGTCGTCGCGCGGCTGAGTACCAACACCTGCTCAGTTGGC 1130
 DB 724 GAGATATGCTGTCGTCGTCGTCGCGCGGCTGAGTACCAACACCTGCTCAGTTGGC 783
 QY 1131 TTCGCGCCCTGAGACCTCTTACAGCTTGGCATGGAGGAGCTGAGCCCGGGAGAGGT 1190
 DB 784 TTCGCGCCCTGAGACCTCTTACAGCTTGGCATGGAGGAGCTGAGCCCGGGAGAGGT 843
 QY 1191 GTGCATTGTAACCCAGCTGCTTACCGCGCTGGGCGTCTCTGTAGAAGAGCGAGC 1250
 DB 844 GTGCATTGTAACCCAGCTGCTTACCGCGCTGGGCGTCTCTGTAGAAGAGCGAGC 903
 QY 1251 TTCACCCACTGGGCGATCTCCGGGCGAGGAGCTGA 1286
 DB 904 TTCACCCACTGGGCGATCTCCGGGCGAGGAGCTGA 939

RESULT 8
 091640
 ID 091640 standard; cDNA; 1056 BP.
 AC 091640;
 DT 14-MAR-1996 (first entry)
 DE Mouse Indian hedgehog protein gene.
 KM Mouse Indian hedgehog gene; probe; primer; diagnostic;
 KW nervous system disorder; gene therapy; antibody; ds.
 OS Mus musculus.
 FH Key Location/Qualifiers
 FT cds 1..1008
 FT /*tag= a
 FT /product= Mouse Indian hedgehog protein
 PN W09518856-A1.
 PD 13-JUL-1995.
 PF 30-DEC-1994; U14992.
 PR 30-DEC-1993; US-176427.
 PR 14-DEC-1994; US-356060.
 PA (HARD) HARVARD COLLEGE.
 PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
 PI Ingham PW, McMahon AP, Tabin CJ;
 DR WPI: 95-255060/33.
 DR P-PSDB: R77343.
 PT Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful
 PT to treat degenerative nervous system disorder(s) and in gene
 PT therapy.
 PS Claim 4; Page 137-38; 210pp; English.
 CC The sequence encodes a mouse Indian hedgehog protein, homologous
 CC to a Drosophila hedgehog protein (R77337), and has been isolated by
 CC low stringency screening of a mouse genome DNA library and
 CC screening of an 8.5 day post coitum cDNA library. The partial cDNA
 CC is complete at the 3'-end, as evidenced by the presence of a
 CC polyadenylation consensus sequence and short poly-A tail. Probes
 CC and primers derived from hedgehog sequences may be used as
 CC diagnostic agents for neuromuscular, autonomic or central nervous
 CC system disorders, and the gene may also be used in gene therapy.
 CC Antibodies generated from the encoded protein may be used as
 CC therapeutic or research reagents.
 SQ Sequence 1056 BP; 204 A; 321 C; 312 G; 219 T;

Query Match 52.0%; Score 843.2; DB 1; Length 1056;
 Best Local Similarity 88.2%; Pred. No. 1.3e-157;
 Matches 928; Conservative 0; Mismatches 123; Indels 1; Gaps 1;

QY 276 GAGCGCTTCAAGAGCTACCCCAATTAATCAGATCATCTTCAAGAGAGAGAG 335
 ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||

Db	1	GAGGGCTTCAAGAGCTCAACCCCACTCAATCCGACATATCTTCAAGACGAGAG	60
Qy	336	AACACAGAGCCCGACCGCCCTCATAGACCAGCGCTGCAGAGACCGCCTAACTCGCTGCT	395
Db	61	AACACGGGTCCCGACCGCCCTCATGACCAGCGCTGCAGAGACCGCCTAACTCGCTGCGC	120
Qy	396	ATCCGGGTGATGAAACAGTAGTGGCCCGGTGTGAAGCTGCGGGGTACCGAGAGGGGTGGGACGAG	455
Db	121	ATCTGTGTATGAACAGTAGTGGCTGGTGTGTAACTGTGGGTATCCGAAAGCGCGGGATGAA	180
Qy	456	GACGGCCACCACTCAGAGAGTCCCTGCATTTATGAAGGCGCGCGCGTGGAGATACACACA	515
Db	181	GATGGCCATCACTCAGAGAGTCTTTACACTATGAGAGGCGCGGGGTGGATATCACACAC	240
Qy	516	TCGAACCCGACGACCAATTAATATGAGACTGTGGCGCGCTTGGCAGTGAAGCGCGCTTT	575
Db	241	TCGAACCGGTACCGAATAATATGATGAGACTGTGGCGCGCTTGTAGCAGTGAAGCGCGCTTC	300
Qy	576	GACGGGTGATTAACGATTAAGGGCCACGTGCATTTGTCCGTAAATCCGACACACG	635
Db	301	GACTGGGTGATTAACGATTCAGAGGCCACAGTGCATTTGCTGTAAAGTGTGAGATTGC	360
Qy	636	GCCCGACACCAAGACGCGCGCGCTGCTTCCCTGCGGAGCCCAAGTACGGCTGGAGATGGG	695
Db	361	GCCCGCTCCCAAGACAGGTGGCTGCTTCTCGCGGAGCCCAAGTACGGCTTGAAGAACGGG	420
Qy	696	GCGCGTGTGGCCTTTGTACGCCGTGAGGCCGGGAGACCGTGTGCTGGCATGGGGAGAT	755
Db	421	GAGCGTGTGGCCTTGTACGCTGTAAAGCCAGAGACCGCGGTCTGGCATGGGGAGAT	480
Qy	756	GGGAGCCCAACCTTCAACGATGTGCTATTTCCTGTGAGACCGGACCCCAAGCTTGA	815
Db	481	GGGAGCCCAACCTTCAATGTGTATTTCCTGTGAGACCGGACCCCAACCGCTGAGA	540
Qy	816	GCTTCCAGGATCGAGATTCAGAGACCCCGCGCGCTGGACATCAACCCGCTAC	875
Db	541	GCTTCCAGGATCGAGATTCAGAGATCTCCGCGCTGGCTGGCTTACGCTGTCCAC	600
Qy	876	CTGCTCTTAAAGGCTGTACAAATCAACAGGAGCCGGACCGCGTTCCGGGACCAATTTGCC	935
Db	601	CTGCTCTTCAATGGGAGCATTCATACAGAACAGCAGCCCATTCGCGGACCAATTTGCC	660
Qy	936	AGCCAGTGCAGGCTGGCCAGTACGTGTGTGTGGCTGGGGTGCACAGGCTTGCAGGCTGCC	995
Db	661	AGCCATGTGCACCGACCAATATGTGTGTATCAAGGGGTACAGGCGCTCCAGGCTGCT	720
Qy	996	GCGGTGGAGCTGTCTTACACAGTGGGCCCTGAGGGCCCTACGCGCGGCTCAAAAGAT	1055
Db	721	GCGGTGGAGCTGTCTTCAACAGTGGGCCCTTGGGCTCTTATGTCTTCTTCAAAAGGAT	780
Qy	1056	GGGACACTGTGTGTGAGAGATGTGTGTGACATCTCGTTCGCGGCGGTGGCTACCCACAC	1115
Db	781	GGGACACTGTGTGTGAGAGATGTGTGTGACATCTCGTTCGCGGCGGTGGCTACCCACAT	840
Qy	1116	CTGGCTCAGTTGGCTTGTGGCCCTTGAGACTCTTTTACAGCTTTGGCATGGGCACTGG	1175
Db	841	CTGGCTCAGTTGGCTTGTGGCCCTTGAGACTCTTTTCCAGTTTGGCATGGGCACTGG	900
Qy	1176	ACCGCGGGGAGGGGTGTGAAATGTGTAATCCCGACACTGCTTACCGCGCTGGCGGCTCGTG	1235
Db	901	ACCGCATGTGGGTGTGTACTCTTACCTTCAATATCTCTTACCGCGCTGGCGGCTCTTGTG	960
Qy	1236	CTAGAAAGAGGACACTTTCACACCACTGGGATGTCCGGGGACAGGAGCTGAAGAGACTC-	1295
Db	961	CTAGAAAGAGACACTTTCATCCACTGTGGGATGTCTGGGGCAAGAGCTGAAGGAGACT	1020
Qy	1295	CACCGCTCCCTCCTGGAATGCTGTACTGGG	1326
Db	1021	AACCACTGCTCCTCTGGAAATGCTGTGCTGGG	1052

ID	Accession	Standard	CDNA	1277	BP
AC	091636	standard	CDNA	1277	BP
DE	091636				
DI	22-FEB-1996	(first entry)			
DE	Chicken	sonic hedgehog protein gene.			
KW	Chicken	sonic hedgehog gene; nested polymerase chain reaction;			
KW	stage 22/22 limb bud;	transgenic mouse screening; probe; primer;			
KW	PCR;	diagnostic; nervous system disorder; gene therapy; antibody;			
KW	ds				
OS	Gallus domesticus.				
PN	W09518856-A1.				
PD	13-JUL-1995.				
PF	30-DEC-1994; U14992.				
PR	30-DEC-1993; US-176427.				
PR	14-DEC-1994; US-356060.				
PA	(HARD) HARVARD COLLEGE.				
PI	(IMCR) IMPERIAL CANCER RES TECHNOLOGY.				
PI	Ingram PW, McMahon AP, Tabin CJ;				
DR	WPI: 95-255060/33.				
DR	P-Psdy; R77338.				
PT	Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful				
PT	to treat degenerative nervous system disorder(s) and in gene				
PT	therapy.				
PS	Claim 2; Page 133-35; 210pp; English.				
CC	The sequence encodes a chicken sonic hedgehog protein, homologous				
CC	to a Drosophila hedgehog protein (R77337), and is isolated by				
CC	nested polymerase chain reaction using primers VH50 (Q91643),				
CC	VH80 (Q91644) and VH31 (Q91645). A clone resulting from the				
CC	amplification (encoding R77348) is used to isolate the full-length				
CC	CDNA from a stage 22/22 limb bud cDNA library. Primer Q91649 is				
CC	derived from the chicken CDNA sequence, and is used to screen				
CC	transgenic mice. Probes and primers derived from the sonic				
CC	hedgehog sequence may be used as diagnostic agents for				
CC	neuro muscular, autonomic or central nervous system disorders, and				
CC	the gene may also be used in gene therapy. Antibodies generated				
CC	from the encoded protein may be used as therapeutic or research				
CC	reagents.				
SQ	Sequence 1277 BP; 262 A; 396 C; 387 G; 232 T;				
Query Match	28.3%; Score 459; DB 1; Length 1277;				
Best Local Similarity	65.2%; Pred. No. 3.9e-82;				
Matches 763:	Conservative 0; Mismatches 360; Indels 48; Gaps				
QY	163	GGCGACCGCCACGCAAACTGTCGCCGCTGCTACAGACGATTCAGCCCAATGTCGCCG	222		
DB	107	GGAGCGACCCCAAAAAGCTGACCCCGTTAGCCTATAGACGATTTATTCCTATGTCGAC	166		
QY	223	AAGAAGCCTGGGGCGCCGACGCGCTATGAAAGGCAAAATGCTCTGGCACTCCGAGCGT	282		
DB	167	ACAAGACCTTAGGGCCCACTGGAAATATGAAAGGGAAGTCAAGAAATCCCGAGAGAT	226		
QY	283	TCAAGAGCTCACCCCAATTACAAATCCAGCATATCTTCAAGGACGAGAGACACAG	342		
DB	227	TTRAAGAACTAACCCCAATTTCAAAACCTCGACATATTATTTTAAAGATGAAGGAACAGG	286		
QY	343	GGCGGACCGCCTCATGACCCAGCGCTGCAAGAGACCGCCTGAATCGCTGCTATCTCGG	402		
DB	287	GAGCTGACGACTGATGACTCAGCGCTGCAAGACCAAGCTGAATGCGCTGGCATCTCGG	346		
QY	403	TGATGAACCAATGCGCCGGTGTGAAGCTGCGGGTGAACCGAAGGGCTTGGACGAGAGCGCC	462		
DB	347	TGATGAACCAATGCGCCGGTGTGAAGCTGCGGGTGAACCGAAGGGCTTGGACGAGAGCGCC	406		
QY	463	ACCACGACAGAGAGTCCCTGCATTTATGAGGGGCGCGCGGGTGGACATCACCACATCGACC	522		
DB	407	ATCACTCCGAGAAATGCTGCTGCACTACGAGAGTGTGCGCGCTGGACATCACCACATCGGATC	466		
QY	523	GGGACCGCAATATGAGACTGTGCGCGCGCTTGGCACTGAGAGCGCGGCTTGAATGGG	582		
DB	467	GGGACCGCAATATGAGAAATGCTGCGCGCGCTGCGCTGAGAGCGCGGCTTGAATGGG	526		
QY	583	TGATTAACAGTCAAGAGCGCCACGTCATTTGCTCGGTCAAGTCCGAGCACTGGCGCGAG	642		

```

Db 527 TCTACTAGATGCCAAGGCGCAATCCACTGCTCCGTCAAGAGAAAACCTAGTGGCAG 586
Qy 643 CCAAGACGGGGGCTGCTTCCCTGCGGAGCCAGGTACGCTGGAGAGGGGCGCTG 702
Db 587 CGAATACAGAGAGGCTGCTTCCCTGCTGACGCCACAGTGCAGCTGGAGCATGGAGGACCA 646
Qy 703 TGGCCTGTACGCGGTAGAGCGGGAGACCGTGTGTGCTGCTGAGGGAGATGGAGCC 762
Db 647 ACCTGTGAAGAACCTGACCCCTGGGGAGCCGGGTGCTGCTGTGACGGAGCGGCGGC 706
Qy 763 CCACCTTCAGCATGTGCTCATTTCTCTGAGCCGAGCCGACAGGCGGAGAGGCTTCC 822
Db 707 TCTCTACAGTGTCTTCTTCCCTTCTTCTGACCGGATGAGACAGCTCCGAAAGCTTCT 766
Qy 823 AGGTACAGAGATCAGAGACCCCGGAGCGGCTGAGCATCACCCTGCTACCTGCTCT 882
Db 767 AGGTATGAGAGAGCGGAGCCCGGCGGCTGCTGCTGAGAGGGGCGGACCTGCTCT 826
Qy 883 TTACGGCTGACATCAGACGAGCGG-----CAGCCGCTTCCGGGCGACAT 930
Db 827 TTGTGGCCCCCAGCAGCAACAGTCCGAGCCAGAGGTCCACAGTGGCCAGGCGCTCT 886
Qy 931 TTGCCAGCCAGTGGAGCGCTGGGCA-----GTAAGTGTGGGTGGGTGGCCAGGCG 984
Db 887 TGGCCAGCAGCGTGAAGCGCTGGCCAGCGTGTCTATGTGCTGGGCGGAGCGGACGAC 946
Qy 985 TGCAGCTGCCCCGCTGGAGCTGTCTTACAC---ACGTGGCCCTCGGGGCTACGCCC 1041
Db 947 TCTGTGCGGCGCTGTCTGACAGAGGTCTCATTTGGGAGAGAGGCTCGGAGGCTACGCC 1006
Qy 1042 CGCTCAGAAAGCATGGGACACTGTGTGTGGAGAGATGTGTGCTGCTTCTTCCGGCCG 1101
Db 1007 CACTCACCCCGCAGGCGACCATCTCTATCACCGGGGTGTGTGGCTCTCTGACGCTCA 1066
Qy 1102 TGGCTACACCAACCTGAGTGGCTTGGCCCTTGGCCCTGAGACTTTCACAGATTGG 1161
Db 1067 TGGAGAGACAGTGTGGGCGCATTTGGGCGCTTGGCACCATTCCGCTTGGCTCAGGGCTCC 1126
Qy 1162 CATGGGCGACGTGGACCCCGGGGAG-----GGTGTGC 1194
Db 1127 TGGCGGCGCTGTGCGCCAGATGGGGGCGCATCCCTACTGCGCAGCAGCAGCAGTGGCATCC 1186
Qy 1195 ATTGGTACCCCAAGCTGCTTACCGCTTGGGCGCTCTCTGCTAGAAAGGCGAGCTTCC 1254
Db 1187 ATTGGTACCTCAGGCGCTCTTACCGCATGCGCATGCGAGCTGGGTGTGTGAGCGGCTGC 1246
Qy 1255 ACCCACTGGGCAATGTCCGGGCGAGGAGCTG 1285
Db 1247 ATCCGCTGGGCAATGTGCGACCGGCGCAGCTG 1277

RESULT 10
X16182
ID X16182 standard; cDNA: 1277 BP.
AC X16182;
DE 29-APR-1999 (first entry)
BT Chicken Shh hedgehog cDNA sequence.
DI Patched; hedgehog; plc therapeutic; neuroprotective; neuronal cell;
KW brain infarction; cerebral infarction; transient ischemic attack;
KW stroke; cerebral infarct volume; spinal cord; oedema; trauma;
KW haemorrhage; encephalomyelitis; coronary bypass; cerebral hypoxia; ds.
OS Gallus sp.
FH key
FT 1. 1275
FT Location/Qualifiers
FT CDS
FT /tag= "a"
FT /product= "hedgehog sequence"
FT /note= "no stop codon given"
PN WO9900117-82.
PD 07-JAN-1999.
PE 26-JUN-1998; U13387.
PR 27-JUN-1997; US-883656.
PA (ONTO-) ONTOGENY INC.
PI Mahanthappa NK.

```

```

DR WPI: 99-095458/08.
DR P-PSDB: W94468.
PT Method for limiting damage to neurons caused by ischemic or epoxic
PT conditions - is used for the treatment and prevention of e.g.
PT cerebral infarction, stroke and transient ischemic attacks
PT Disclosure: Page 63-64, 104pp. English.
CC A method has been developed for limiting the damage to neuronal cells by
CC ischemic or epoxic conditions by administering a plc (patched)
CC therapeutic agent to reduce cerebral infarct volume (CIV). Damage to
CC neuronal cells can also be limited by administering a gene activation
CC construct which recombines with the genomic hedgehog gene to provide a
CC heterologous transcription regulator linked to the coding region of this
CC gene. Administration of the plc therapeutic agent is used to protect
CC cerebral tissues against ischemic injury; to treat cerebral infarct or
CC ischemia, stroke (thrombotic or embolic) and transient ischemic
CC attacks. It may also be used as a prophylactic in many other cases of
CC injury to the brain or spinal cord, oedema caused by trauma, hemorrhage
CC and encephalomyelitis, or in conjunction with (coronary bypass) surgery.
CC treatment (which may be prophylactic) is used where ischemic/epoxic
CC conditions may cause cerebral hypoxia, or progressive loss of neurons
CC due to oxygen depletion, including in patients with hypotension. The
CC treatment reduces CIV by at least 25, particularly at least 70%. The
CC present sequence encodes a hedgehog sequence given in the present
CC invention.
SQ Sequence 1277 BP: 262 A: 396 C: 387 G: 232 T:

Query Match 28.3%; Score 459; DB 1: Length 1277;
Best Local Similarity 65.2%; Pred. No. 3.9e-82;
Matches 763; Conservative 0; Mismatches 360; Indels 48; Gaps 4;

Qy 163 GCGACCGCCACGCAAACTGTCGCTGCGCTGCTCAAGCACTTACGCCCAATGTGCCG 222
Db 107 GGAGGACACCCCAAAAGCTGACCCGTTAGCTTAAACAGATTATTCCTCAATGTGGCAG 166
Qy 223 AGAAGACCTTGCGGCGCCAGCGGAGCGCTTGAAGCAAGATGCTCGACGCTCCAGCGCT 282
Db 167 AGAAGACCTTAGGGGCGCTGAGATGAGATGAGAGAAATACAGAAAGCTCCGAGAT 226
Qy 283 TCAAGAGCTCACCCCAATTACATCCAGACATCATCTTCAAGAGAGAGAAACAG 342
Db 227 TTAAAGACTAAACCCCAATTACAACTGACATTAATTTTAAGATGAAGAAACAG 286
Qy 343 GCGCGAGCGCTCATGACCCAGCGGCTGCAAGAGACCGGCTGAATCGCTGCTATCTCG 402
Db 287 GAGCTGACAGACTGATGACTGACGCTGCAAGAGCAACCTGAATCCCTGCGCATCTCG 346
Qy 403 TGATTAACAGAGTGGCGCGGTGTAAGCTGCGGCTGACCGAGGCGTGGAGACAGAGCGCC 462
Db 347 TGATTAACAGAGTGGCGCGGTGTAAGCTGCGGCTGACCGAGGCGTGGAGAGATGGCC 406
Qy 463 ACCACTCAGAGAGTCCCTGATTTAGAGGCGCGCGGTGGAGATCCACCATACAGCC 522
Db 407 ATCACTCCGAGAGATCGCTGACATCGAGGGTCCGCGGTGGAGATCCACAGTGGATC 466
Qy 523 GCGACCGCAATATGATGACTGTGCGCGCTTGGCAGTGGAGGCGGCTTTGACTGG 582
Db 467 GGGACCGGAGCAAGTACGAATGTGCGCGCTGCGCTGCGAGGCGCGCTTGCAGTGG 526
Qy 583 TGTATTAGAGTCAAAAGGCCACAGTGTGCTGCTGTAAGTCCAGAGACTCGCGCGAG 642
Db 527 TCTACTACAGTCCAAAGCGCACATCCCTGCTGTAAGAGCAAGAAATCAGTGGGAG 586
Qy 643 CCAAGACGGGCGCTGCTTCCCTGCGGAGCCAGGTACGCTGGAGAGTGGGCGCGCTG 702
Db 587 CGAATACAGAGAGCTGCTTCCCTGCTGACGCCACAGTGCAGCTGAGAGATGGAGGACCA 646
Qy 703 TGGCCTGTACGCGGTAGAGCGGGAGACCGTGTGTGCTGCTGAGGGAGATGGAGCC 762
Db 647 AGCTGTGAAGAGCTGAGCCCTGGGAGACCGGCTGCTGCTGAGACGAGCGGCGGC 706
Qy 763 CCACCTTCAGCATGTGCTCATTTCTCTGAGCCGAGCCGACAGGCTGAGAGCTTCC 822

```


Db 887 TCGCCAGCACTGAAAGCTGGCCAACTGTCTATGTCTGGGCGGAGGCGGCACTAC 946
 Qy 985 TCGAGCTCCCGCGTGGAGCTGTCTACAC---ACGTGGCCCTCGGGGCGCTACGCC 1041
 Db 947 TCGTGGCGGCGTGTCTCCACAGAGTCTCATGTGGGAGAGAGGCGTCCGAGCTACGCC 1006
 Qy 1042 CGCTACAAAGATGGGACACTGTGTGTGGAGATGTGTGCTCTCTCTCGGCGCG 1101
 Db 1007 CACTACACCGCCAGGCGACCATCTCATACACCGGCTGTGGCTCTCTCTCGGCTCA 1066
 Qy 1102 TGGCTACACCACTGCGTCACTGTGGCTGTGGCGCCCTGAGACTCTTCAAGCTTG 1161
 Db 1067 TCGAGAGACACAGTGTGGGCGCATGTGGCGCTTGGCACCATCTCGCTGCTCAGGGCTGC 1126
 Qy 1162 CATGGGCGACGTGAGCCCGGGGGAG-----GGTGTGC 1194
 Db 1127 TGGCGCCCTCTGCCAGATGGGGCCATCCCTACTGCGCCACACACCACTGGGATCC 1186
 Qy 1195 ATTGTACCCCAAGCTGTCTACCGCTGGGCGCTCTCTCTGTAAGAGGCGAGCTTC 1254
 Db 1187 ATTGTACACAGGCGCTCTCTACCGCATGTGGCACTGGGCTGTGTGATGTGACGGCTGC 1246
 Qy 1255 ACCCACTGGGCACTGTCGGGGGAGAGGCTG 1285
 Db 1247 ATCCGCTGGGCACTGTGTGCGACCGGCGAGCTG 1277
 RESULT 12
 X25098
 ID X25098 standard; cDNA; 1277 BP.
 AC X25098:
 DI 05-JUL-1999 (first entry)
 DE Chicken Sonic hedgehog protein Shh cDNA.
 KW Sonic hedgehog; Shh gene; chicken; hedgehog therapeutic;
 KM ptc therapeutic; patched; chicken transduction; muscle atrophy;
 KW cachexia; muscular myopathy; myoblastic sarcoma; therapy; ds.
 OS Gallus sp.
 PN W09J10004.A2.
 PD 04-MAR-1999.
 PR 28-AUG-1998; 017922.
 PR 29-AUG-1997; US-057394.
 PA (ONTO-) ONTOGENY INC.
 PI Bladen CS, Currie PD, Hughes SM, Ingham PW;
 DR WPT; 99-243557/20.
 P-PSDB: Y05510.
 PS A new method to regulate muscle growth
 PT Disclosure; Page 92-94; 130P; English.
 CC This nucleotide sequence comprises a coding region for the chicken
 CC Sonic hedgehog protein Shh (see Y05510). The invention relates to
 CC a method for modulating the formation and/or maintenance of muscle
 CC tissue by ecotopically contacting muscle cells, especially muscle
 CC stem/progenitor cells, in vitro or in vivo, with a hedgehog
 CC therapeutic (i.e. hedgehog polypeptides and gene therapy
 CC constructs) or ptc therapeutic (i.e. a small organic molecule that
 CC mimics the effect of hedgehog proteins on patched signalling) or
 CC activates or potentiates patched signalling) in an amount effective
 CC to alter the growth state of the treated cells. Also claimed is a
 CC method for treatment or prevention of disorders of, or surgical or
 CC cosmetic repair of, such muscle tissues, by administering a
 CC hedgehog polypeptide or ptc therapeutic. The disorder may be
 CC muscle atrophy, in particular skeletal muscle atrophy or cardiac
 CC muscle atrophy, cachexia, or muscular myopathy (all claimed). The
 CC hedgehog polypeptide or ptc therapeutic can inhibit growth of
 CC myoblastic-derived tissue to provide treatment of hyperplastic or
 CC neoplastic growth of muscle tissue such as in myoblastic sarcoma
 CC (also claimed). The hedgehog therapeutic preferably comprises at
 CC least a bioactive extracellular portion of a hedgehog protein (see
 CC Y05510-19) encoded by a vertebrate hedgehog gene (see X25098-107),
 CC especially a human hedgehog gene.
 SQ Sequence 1277 BP; 262 A; 396 C; 387 G; 232 T;

Query Match

28.3%; Score 459; DB 1; Length 1277;

Best Local Similarity 65.2%; Pred. No. 3,9e-82;
 Matches 763; Conservative 0; Mismatches 360; Indels 48; Gaps 4;
 Qy 163 GAGCGCCGCCAGCAAACTCGTCCGCTGCGCTTCAAGCACTTCAAGCCATGTGCCG 222
 Db 107 GGAGGACCCCAAAAGCTGACGCCCGTTAGCTTAAGCACTTTATTCCAATGTGGAG 166
 Qy 223 AAGAGACCTTGGGCGCCAGCGGAGCGTATGAGGCAAGATGCTTCGAGTCCGAGCGT 282
 Db 167 AAGAGACCTTGGGCGCGAGTGAAGATATGAAGGGAATATCAAGAAACTCCAGAGAT 226
 Qy 283 TCAAGAGCTCAACCCCAATTCATTCAGACATCATCTTCAAGAGACAGAGAAACAG 342
 Db 227 TTAAGAACTAACCACCAATTCACACCTTGACATTTATTTAAGAGATTAAGAAACAG 286
 Qy 343 GCGCGGACGCGCTCATGACCGAGCGCTGCAAGACCGCGCTGAGCTGTGCTATCTCG 402
 Db 287 GAGCTGACAGACTGATGACTAGCGCGTCAAGAGCAAGCTGATATGCTTGGCGATCG 346
 Qy 403 TGATGAACCACTGGCCGCGTGTGAAGCTGCGGGTGAACCGAGGGCTGGAGACGAGCGC 462
 Db 347 TGATGAACCACTGGCCGCGGTTGAAGCTGCGGGTGAACCGAGGGCTGGAGAGGATGCGC 406
 Qy 463 ACAAATCAGAGAGAGTCCCTGATATATGAGGCGCGCGGTGGAATATCAATCAAGAC 522
 Db 407 ATCACTCCGAGAGATCGGTGCACTACGAGGGTGGCGCGCTGGATCAACAGTGGATC 466
 Qy 523 GCGACCGCAATAGATGATGACTGTGGCGCGCTTGGAGTGAAGCGCGCTTGTACTGG 582
 Db 467 GGGACCGGAGCAATGAGGATGTGTGGCGCGCTTGGAGTGAAGCGCGCTTGTACTGG 526
 Qy 583 TGTATTACGAGTCAAAAGGCCCAAGTGTCTCTCGTCAAGTCCGAGCACTCGGCGGAG 642
 Db 527 TCTATTAGAGATCCAAAGCGGCAATCTCACTGCTCCGTAAGAGAAACTCAGTGGAG 586
 Qy 643 CCAAGACGGGGCGCTTCCCTCGCGGAGCCAGTACGCTGGATGAGTGGGCGCGCTG 702
 Db 587 CGAAATCAGAGAGGCTGCTCCCTGCTGCTGAGCAAGTCACTGAGATGAGAGGACCA 646
 Qy 703 TGGCTTGTACCGCTGTAGGCGCGGAGACCGTGTCTGCTGCGCATGGGAGGATGGAGCC 762
 Db 647 AGCTGTGAAGAGCACTGAGCCCTGGGGAGCCGCGTGTGTGCTGCTGAGCGAGCGCGC 706
 Qy 763 CCACCTTACGAGATGTCTCATTTTCTGTAACCGGAGCCCAAGAGCTGAGAGCTTCC 822
 Db 707 TGCTTACAGTACTCTCTCACTTCTCGACCGGATGAGAGCTCCCGAAGCTTCTCT 766
 Qy 823 AGGTATGAGAGTCAAGAGACCCCGGCGCGCTGCGCATCAACCGCGCTCACTGCTCT 882
 Db 767 AGGTATGAGAGCGGAGAGCCCGGCGCGCTGTACTAGACGGCGGCCACCTGCTCT 826
 Qy 883 TTACGGCTGCAATATCAACGAGAGCGG-----CAGCCGCTTCCGGGCCACAT 930
 Db 827 TTGTGGCGCCCGAGACACACAGTGTGAGGCGACAGGCTCCACAGTATGCGAGGCGCT 886
 Qy 931 TTGCGACCACTGACGCTTGCA-----GTACGTCTGTGTGGTGGTGGTGGCCAGGCC 984
 Db 887 TCGCAGCAAGTGAAGGCTGCGCAACGTGTATGTCTGGGCGAGGGCGGAGCAGC 946
 Qy 985 TCGACCTTCCCGCGGTGGCACTGTCTTACAC---ACGTGGCGCTGGGGGCTTACGCC 1041
 Db 947 TGTCTCCGCGCTGTGTCCACAGCGCTTATGTGGGAGAGGCTCGGAGCTTACGCC 1006
 Qy 1042 CGCTACAAAGATGGGACACTGTGTGTGGAGATGTGTGTCATCTGCTCGCGGCGG 1101
 Db 1007 CACTACACCGCCAGGCGACCATCTCATACACCGGGTGTGTGCTCTCTCTCAACCGCGCA 1066
 Qy 1102 TGGCTGACCAACACTGCTGATGTGGCTTGTGGCCCTGAGACTTTTACAGCTTGG 1161
 Db 1067 TCGAGAGACACAGTGTGGCGCATTTGGGCTTGCACACCATCTGCTGCTGAGGGGCTGC 1126
 Qy 1162 CATGGGCGACGTGAGCCCGGGGGAG-----GGTGTGC 1194

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2000, 07:55:21 : Search time 975.44 Seconds
(without alignments)
4948.935 Million cell updates/sec

Title: US-08-900-220-8

Perfect score: 1191
Sequence: 1 ATGGCTCTCTGACCAATCT.....CGAGAGACTACTGGGCTGA 1191

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: em_est20:*
21: em_est21:*
22: em_est22:*
23: em_est23:*
24: em_est24:*
25: em_est25:*
26: em_est26:*
27: em_est27:*
28: em_est28:*
29: em_est29:*
30: em_est30:*
31: em_est31:*
32: em_est32:*
33: em_est33:*
34: em_est34:*
35: em_est35:*
36: em_est36:*
37: em_est37:*
38: em_est38:*
39: em_est39:*
40: em_est40:*
41: em_est41:*
42: em_est42:*
43: em_est43:*
44: em_est44:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: gb_est33:*
53: gb_est34:*
54: gb_est35:*
55: gb_est36:*
56: gb_est37:*
57: gb_est38:*
58: gb_est39:*
59: gb_est40:*
60: gb_est41:*
61: gb_est42:*
62: gb_est43:*
63: gb_est44:*
64: gb_est45:*
65: gb_est46:*
66: gb_est47:*
67: gb_est48:*
68: gb_est49:*
69: gb_est50:*
70: gb_est51:*
71: gb_est52:*
72: gb_est53:*
73: gb_est54:*
74: gb_est55:*
75: gb_est56:*
76: gb_est57:*
77: gb_est58:*
78: gb_est59:*
79: gb_est60:*
80: gb_est61:*
81: gb_est62:*
82: gb_est63:*
83: gb_est64:*
84: gb_est65:*
85: gb_est66:*
86: gb_est67:*
87: gb_est68:*
88: gb_est69:*
89: gb_est70:*
90: gb_est71:*
91: gb_est72:*
92: gb_est73:*
93: gb_est74:*
94: gb_est75:*
95: gb_est76:*
96: gb_est77:*
97: gb_est78:*
98: gb_est79:*
99: gb_est80:*
100: gb_est81:*
101: gb_est82:*
102: gb_est83:*
103: gb_est84:*
104: gb_est85:*
105: gb_est86:*
106: gb_est87:*
107: gb_est88:*
108: gb_est89:*
109: gb_est90:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result

Query

No.	Score	Match length	DB	ID	Description
1	254	21.3	314	28	AA064660 zml3f08.s
2	239	20.1	337	49	AI645932 mml1g06.y.y
3	234.8	19.7	463	49	AI666359 mml1g06.x
4	171.6	14.5	299	69	AM144802 EST291875
5	121.8	10.2	406	63	AI958076 fcg9e01.y
6	72.6	6.1	932	82	AI066742 cef70h05.s
7	71	6.0	459	41	AI015155 cef70h05.s
8	70.4	5.9	1100	83	AI106855 Drosophila11
9	70.2	5.9	925	82	AI053013 Drosophila11
10	67.8	5.7	512	30	AA245525 my52c03.x
11	67	5.6	432	27	AI325370 m134b11.y
12	65.4	5.5	354	24	AA032692 m134b11.x
13	63.8	5.4	935	85	AI066501 Drosophila11
14	63	5.3	1101	83	AI084460 Drosophila11
15	62	5.2	932	82	AI066742 Drosophila11
16	61.6	5.2	935	82	AI053013 Drosophila11
17	61.6	5.2	935	82	AI053013 Drosophila11
18	59.8	5.0	776	82	AI099333 Drosophila11
19	58.8	4.9	839	82	AI054280 Drosophila11
20	56.6	4.8	1101	82	AI076875 Drosophila11
21	56	4.7	1203	83	AI06054 Drosophila11
22	55.4	4.7	970	82	AI098787 Drosophila11
23	55.2	4.6	1203	83	AI106054 Drosophila11
24	54.4	4.6	1201	83	AI106648 Drosophila11
25	53.6	4.5	915	83	AI106083 Drosophila11
26	53	4.5	991	74	AF122121 AF122121
27	52.2	4.4	1101	83	AI104440 Drosophila11
28	51.8	4.3	977	82	AI076850 Drosophila11
29	51.8	4.3	1009	82	AI098682 Drosophila11
30	51.6	4.3	1009	82	AI098682 Drosophila11
31	51.6	4.3	1101	83	AI104949 Drosophila11
32	51.2	4.3	910	82	AI065629 Drosophila11
33	50.8	4.3	955	82	AI065794 Drosophila11
34	50.4	4.2	1101	82	AI063912 Drosophila11
35	49.8	4.2	645	82	AI015159 Drosophila11
36	49.2	4.1	1036	82	AI096770 Drosophila11
37	49	4.1	1077	83	AI105068 Drosophila11
38	48.9	4.1	1101	83	AI108768 Drosophila11
39	48.6	4.1	692	82	AI050923 Drosophila11
40	48.6	4.1	844	82	AI050565 Drosophila11
41	48.6	4.1	962	82	AI077188 Drosophila11
42	48.6	4.1	1101	83	AI106750 Drosophila11
43	48.4	4.1	1171	84	B10823 r11e7-77
44	48.2	4.0	918	91	AI0895329 HS_4832_A
45	48	4.0	834	82	AI078726 Drosophila11

ALIGNMENTS

RESULT	1
AA064660	
LOCUS	
DEFINITION	314 bp mRNA EST 23-DEC-1997
ACCESSION	U033708.1
VERSION	1
KEYWORDS	
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS	Eutheria; Primates; Catarrhini; Homiidae; Homo.
	1 (bases 1 to 314)
	Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiappelli, B.,
	Chisoe, S., Dietrich, N., Dubugue, T., Favello, A., Gish, W.,
	Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N.,
	Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
	Rohlfing, T., Schellendberg, K., Soares, M.B., Tan, F., Tillery-Meg, J.,
	Trevaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.
	and Marra, M.

FEATURES	SOURCE
<p> BASE COUNT 31 a 100 c 117 g 55 t 11 others ORIGIN </p>	<p> This clone is available royalty-free through LINT ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Possible reversed clone: similarity on wrong strand Possible reversed clone: polyt not found Insert Length: 862 Std Error: 0.00 Seq primer: -40M13 fwd. from Amersham High quality sequence stop: 283. Location/Qualifiers 1..314 /organism="Homo sapiens" /db_xref="GDB:3916972" /db_xref="taxon:9606" /clone="IMAGE:525543" /clone_id="Stratagene pancreas (#937208)" /lab_host="SOBR cells (kanamycin resistant)" /note="Organ: pancreas; Vector: pBluescript SK-; Site.1: EcoRI; Site.2: XhoI; Cloned unidirectionally. Primer: 0190 dt. Pancreatic adenocarcinoma cell line. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGACACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'" </p>
<p> Query Match 21.3%; Score 254; DB 28; Length 314; Best Local Similarity 91.3%; Pred. No. 2,8e-40; Matches 285; Conservative 0; Mismatches 25; Indels 2; Gaps 2; </p>	<p> On May 9, 1995 this sequence version replaced g1:802645. JOURNAL Genome Res. 6 (9), 807-828 (1996) COMMENT 97044478 Contact: Wilson RK Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu WARNING: There is evidence that suggests that the 384-well parent plate of this clone contains both human and mouse derived clones. Thus, the origin of this clone is uncertain. This caution should be kept in mind should you use this clone. </p>
<p> RESULT 2 LOCUS A1645932 537 bp mRNA EST 29-APR-1999 DEFINITION mulipg06.v1 Soares_thymus_2nbmt Mus musculus cDNA clone IMAGE:639130 </p>	<p> Db 299 CTGAGATGCATT 310 </p>

mRNA sequence.
 similar to TR:061488 061488 DESERT HEDGEHOG HOMOLOG PRECURSOR,
 accession
 A1645932
 version
 A1645932.1 GI:4724407
 keywords
 EST.
 source
 house mouse.
 organism
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Scuriognathli; Muridae; Murinae; Mus.
 reference
 1 (bases 1 to 537)
 authors
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 title
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 tumor gene index
 journal
 unpublished (1997)
 comment
 On May 18, 1998 this sequence version replaced gi:3137802.
 contact: Robert Strausberg, Ph.D.
 tel: (301) 496-1550
 email: Robert.Strausberg@nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (infoimage.jhl.gov) for further information.
 This read is a RESEQUENCE of a previously sequenced mouse clone
 This read has been verified (found to hit its original self in the
 correct orientation)
 Possible reversed clone: similarity on wrong strand
 Seq primer: ~40RP from Gibco
 High quality sequence stop: 478.
 location/Qualifiers
 1..537

Query Match	Best Local Similarity	Matches 263: Conservative	20.1%: Score 239; DB 49; Length 537; 86.8%: Pred. No. 2.6e-37; 0: Mismatches 40; Indels 0; Gaps 0;
OY	1	ATGGGCTCTCGACCAACTACTGCGCTTGATGCTGTTGGACATCTGCGCGTGCAGACC	60
Db	235	ATGGCTCTGCGCGCCAGTCTGTGGCCCTGTCTGCTTGGCACTCTGGCACTATCTGCC	294
OY	61	CAGAGCTCGGGGCCGGGGCCGGGGCGGTGGCCGGCGCGCTATGCGCGAAGAGCTC	120
Db	295	CAGAGCTCGGGGCCGGGGCCAGAGACCGGTTGGCCGGCGCGCTATGTCGCAAGCAACT	354
OY	121	GTGGCGGCTACTCTATAAGCAATTTGTGCCCGCGCGTGCACAGACGACCTGGGGCCCACT	180
Db	355	GTGGCTCTGCTATACAAAGCAGTTTGTGCCCATATATGCCAGCGACCTGTGGCGGAGT	414
OY	181	GGGCGCAGCGGAGGAGGCTGCGCAAGGGGCTCCGAGCGCTTCGCGGACCTGTCGCCAAC	240
Db	415	GGGCGCAGCGGAGGAGGCTTACAAAGGGGCTGGAGCGCTTCGCGGACCTGTAACCAAC	474
OY	241	TACAAACCCGACATCACTTCAAGATGAGAGAAACATGAGACCGACCGCCTGATGACC	300

Db 475 TACAAACCCGACAAATCTTCAAGATGAGAGAGAACAACGCGCAGACCGCTGATGACA 534

QY 301 GAG 303
|||

Db 535 GAG 537

RESULT 3
A1666359/c

LOCUS

DEFINITION
A1666359 463 bp mRNA EST 12-MAY-1999
mull1906.x1 Soares;thymus.2NB8T Mus musculus cDNA clone IMAGE:659130
3' similar to YR:061488 061488 DESERT HEDDEHOG HOMOLOG PRECURSOR ;,
mRNA sequence.

ACCESSION
A1666359

VERSION
A1666359.1 GI:4804713

KEYWORDS
EST.

SOURCE
house mouse.

ORGANISM
Mus musculus

REFERENCE
Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Mammalia;
Euthalia; Rodentia; Sciurognath1; Muridae; Murinae; Mus.
1 (bases 1 to 463)
NCI-CCAG <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On May 10, 1998 this sequence version replaced gi:3136856.

COMMENT

JOURNAL

```

Other_Estis: nulligub.yi
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
This clone was previously sequenced on the 5' end only, this new
data is from the 3' end
Possible reversed clone: similarity on wrong strand
High quality sequence stop: 261.

FEATURES
    source
        1..463
            /organism="Mus musculus"
            /strain="C57BL/6J"
            /db_xref="taxon:10090"
            /clone="IMAGE:639130"
            /clone_id="Soares_thymus_2NBMT"
            /sex="male"
            /tissue_type="Thymus"
            /dev_stage="4 weeks"
            /lab_host="DH10B"
            /note="Vector: pT7AD-Pac (Pharmacia) with a modified
polylinker. Site.1: Not I; Site.2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer 15',
TGTATACCATCTCAAGCGAGGCGCGCGCTTTTCTTTTCTTTTCTTTTCTTTT
3'); double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7A3 vector. RNA
provided by Dr. Bertrand Jordan. Library went through two
rounds of normalization, and was constructed by Bento
Soares and M.Fatima Bonaldo."

BASE COUNT      87 a      124 c      138 g      113 t      1 others
ORIGIN
Query Match      19.7%; Score 234.8; DB 49; Length 463;
Best Local Similarity 85.8%; Pred. No. 1.6e-36;
Matches 260; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
0Y      1 ATGGCTCTCTGACCAATCTACTGCGCTTGTCGCTTGACACTTCTGCGCTGCCAGCC 60
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      463 ATGGCTCTGCGGGGCGAGCTGTGTGCGCCCTGTGCTCTTGACACTTCTTGCGCTATCTGCC 404
Matches 260; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
0Y      61 CAGAGCTCGGGGCGGGGCGGGGCGGGTGTGGCCGGCCCGCTATGCGCCGCAAGAGGCTC 120
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      403 CAGAGCTCGGGGCGGGGCGGAGAACCGGGGTGGCCGCGCGCTAATGTCGCGCAAGCAACTT 344

```

```

QY 121 GTCCGCTACTACAGCAATTGTGCCCCGCGTCCACAGCGGACCCCTGGCGCCAGT 180
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 343 GTCCCTCTGTAACAGCAGGTTGTGCCAGTATGCCCCAGCGGACCCCTGGCGCGAGT 284
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 GGGCCAGCGGAGGGGGGGGCGGCAAGGGGCTCCGAGCGCTCCGGGACCCGCGGCCAAC 240
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 283 GGGCCAGCGGAGGGGGGGGCGGCAAGGGGCTCCGAGCGCTCCGGGACCCGCGGCCAAC 224
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 TACACCCCGACATCTCTTCAAGAGTGAAGAGACAGTGGAGCCGACCGCTGATGACC 300
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 223 TACACCCCGACATCTCTTCAAGAGTGAAGAGACAGCGGCGCCGCGCTGATGATCA 164
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 301 GAG 303
    |||
DB 163 GAG 161

RESULT 4
AW144802 299 bp mRNA EST 30-OCT-1999
LOCUS EST291875 Normalized rat embryo, Bento Soares Rattus sp. cDNA clone
DEFINITION RIGIG52.5' end similar to Indian hedgehog protein, mRNA sequence.
ACCESSION AW144802
VERSION AW144802.1 GI:6161619
KEYWORDS EST.
SOURCE Rattus sp.
ORGANISM Rattus sp.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 299)
Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,
Kerlavage,A.R. and Adams,M.D.
Rat Genome Project: Generation of a Rat EST (REEST) Catalog & Rat
Gene Index
Unpublished (1998)
JOURNAL On Jun 5, 1998 this sequence version replaced gi:3189461.
COMMENT Contact: Lee, NH
ATCC The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
For clone availability, additional sequence and expression
information related to this EST please check the TIGR Rat Gene
Index (http://www.tigr.org/tdb/rat/rat.html). To order a clone
contact the ATCC (http://www.atcc.org/atcc.html).
Seq primer: M13 Reverse
FEATURES
Location/Qualifiers
1..299
/organism="Rattus sp."
/db_xref="taxon:10118"
/clone_lib="Normalized rat embryo, Bento Soares"
/dev_stage="embryo 8, 12, 18 dpc"
/note="Vector: pVT3Pac; Site_1: EcoRI; Site_2: NotI"
BASE COUNT 78 a 92 c 80 g 49 t
ORIGIN
Query Match 14.5%; Score 172.6; DB 69; Length 299;
Best Local Similarity 73.6%; Pred. No. 1.7e-24;
Matches 220; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 177 CAGTGGCCAGCGGAGGGGCTGGCAGGGGCTCCGAGCGCTCCGGAGCTCGTGC 236
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 CAGCGGCGCTACTAAGCAAATCGCGCGCAGCTCTGAGCGCTTCAAGGAGCTCACCC 60
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 237 CAACTACAACCCCGACATCTCTTCAAGAGTGAAGAGACAGTGGAGCCGACCGCTGAT 296
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 CAACTACAACCCCGACATCTCTTCAAGAGTGAAGAGACAGTGGAGCCGACCGCTCAT 120
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 297 GACCGAGCGTTCGAGAGAGAGGTGAACGCTTGGCCATTGCCGTATGAACATGTGGCC 356
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

DB 121 GACCCAGCCCTCAAGAGACCGCTCTAATCACTAGCCATCTCTGATGAACCAATGGCC 180
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 357 CGAGAGCGCCCTACAGAGTACTAGAGGCGTGGAGACGAGCGGCGACACCGCTCAGATTTC 416
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 181 GGGTGTGAAGCTGCGGGTATACCGAAGGCTGGGATGAAGACCCCACTACTCAGAGGAATC 240
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 417 ACTCCACTAGAGAGCGCGCTTTGGACATCACTAGCTGACCGGACCGCAACAAGT 475
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 241 TTACACTATGAGGCGCGCGGTGGATATCACACCTCAGACCGGCAAAATAGT 299
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
AI958076 406 bp mRNA EST 20-AUG-1999
LOCUS IC90E01.y1 zebrafish Mashu MPING EST Danio rerio cDNA 5' similar to
DEFINITION TR:092008 092008 VHH-1 PRECURSOR ;, mRNA sequence.
ACCESSION AI958076
VERSION AI958076.1 GI:5750785
KEYWORDS EST.
SOURCE zebrafish.
ORGANISM Danio rerio
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
AUTHORS Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
Cypripnoidea; Cyprinidae; Rasbortinae; Danio.
1 (bases 1 to 406)
Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M.,
Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schuk,R.,
Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.
Mashu Zebrafish EST Project 1998
Unpublished (1998)
JOURNAL On Jun 5, 1998 this sequence version replaced gi:3188836.
COMMENT Contact: Stephen L. Johnson
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrfish@watson.wustl.edu
CDNA Library Preparation: Matthew Clark. cDNA library Arrayed by:
Matthew Clark. DNA Sequencing by: Washington University Genome
Sequencing Center Clone distribution: Genome Systems, St. Louis,
Missouri (web address: www.genomesystems.com) (email contact:
info@genomesystems.com) and Research Genetics, Huntsville, Alabama
(web address: www.resgen.com) (email contact: info@resgen.com) and
ResourcenZentrumPrimateDatenbank, Berlin, Germany (web address:
www.rzpd.de)
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: T3 ET from Amersham
High quality sequence stop: 1.
FEATURES
Location/Qualifiers
1..406
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone_lib="zebrafish Mashu MPING EST"
/dev_stage="XLA-blue MRF"
/lab_host="XLA-blue MRF"
/note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; 1st
strand cDNA was primed with a Not I - oligo(dT)15 primer
[5'GACATAGTCTAGATACCGGACCGCGCGCTTTTCTTTTCTTTT3'];
double-stranded cDNA was ligated to Sal I adaptors (BRL),
digested with Not I and cloned into the Not I and Sal I
sites of the pSPORT1 vector (BRL). Library was constructed
by Matthew Clark (Lehrach lab; ICRF, London and Max Planck
Institut fuer Molekulare Genetik Berlin). cDNAs for EST
analysis were selected following oligonucleotide
hybridization fingerprinting of arrayed clones from
zebrafish late somitogenesis (26 ss), adult liver or

```



```
FEATURES
SOURCE
    Location/Qualifiers
      1..459
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone_image="IMAGE:1622169"
        /clone_lib="Soares total_fetus_Nb2HF8_9w"
        /dev_stage="8-9 weeks"
        /lab_host="DH10B"
        /note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAGTGGAGCGCCCTATTCTTTTCTTTTCTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      106 a      127 c      133 g      93 t
ORIGIN

Query Match      6.0%; Score 71; DB 41; Length 459;
Best Local Similarity 67.1%; Pred. No. 0.0001;
Matches 116; Conservative 0; Mismatches 55; Indels 2; Gaps 1;

OY 82 GGGCCGGTGGCGCGCGCGCTATGCGCGCA--GCAGCTGTCGCGCTACTTACAGC 139
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 454 GGGTGGGGTGGTGGGACCGCGCGCGCGCAAGCACTGTCGCGCTGCGCTACAGC 395
OY 140 AATTGTGCGCGCGCGCGCGAGAGCGACCTGCGCGCGCGCGCGAGGAGG 199
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 394 AATTAGCCCCCATGTCGCGGAGAAACCTGCGCGCGCGCGCTATGAGGAGAG 335
OY 200 TGGCAAGGGGCTCCGAGCGCTTCCGGACCTGCGCGCGCGCGCGAGCCGAC 252
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 334 TGGCTCGCAGCTCCGAGCGCTTCAATGACTACCCCAATTACATCCAGC 282

RESULT 8
CNS016KD      1100 bp      DNA      GSS      26-JUN-1999
LOCUS      Drosophila melanogaster genome survey sequence SP6 end of BAC
DEFINITION      BAC16D22 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION      AL106855
VERSION      AL106855.1 GI:5624152
KEYWORDS      GSS.
SOURCE      fruit fly.
ORGANISM      Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1 (bases 1 to 1100)
AUTHORS      Genoscope.
TITLE      Direct Submission
JOURNAL      Submitted (23-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT      Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (BDGP) -
http://www.edgp.ebi.ac.uk . This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
Project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

FEATURES
SOURCE
    Location/Qualifiers
      1..1100
        /organism="Drosophila melanogaster"
        /plasmid="pBelobAC11"
        /db_xref="taxon:7227"
        /clone_lib="DrosBAC"

BASE COUNT      132 a      229 c      106 g      220 t      413 others
ORIGIN

Query Match      5.9%; Score 70.4; DB 83; Length 1100;
Best Local Similarity 24.4%; Pred. No. 0.00016;
Matches 126; Conservative 175; Mismatches 213; Indels 2; Gaps 1;

OY 616 GGGCCCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 675
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 566 GTGGGSGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCG 645
OY 676 TTGGCGCGATGCTGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 735
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 646 GSGSGGGYKSKSGSGCGGSGGSGGSGGSGGSGGSGGSGGSGGSGGSGGSG 705
OY 736 GACTTGCAGCGCGCGGCTTCAATTGTGCTGTGAGACGAGTGTGCTCAGCA 795
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 706 YCBGSGSSCKCCSBGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 765
OY 766 TTGCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 855
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 766 CTGCTCCCTCTCTCKSKCGCGTGTGCTCKCKCTGTGCTCKBYCTCTT--CRY 823
OY 856 TTGCGACCGGTGTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 915
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 824 BCKGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 883
OY 916 GATGCGCTTGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 975
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 884 KYKCKBCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCK 943
OY 976 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1035
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 944 CCKBCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCK 1003
OY 1036 GTTCTGAGAGCTACCACTGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1095
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1004 CBKCKBCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCK 1063
OY 1096 CTAGGGCGCTGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1131
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1064 BCBCKCKBCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCK 1099

RESULT 9
CNS0091P      925 bp      DNA      GSS      03-JUN-1999
LOCUS      Drosophila melanogaster genome survey sequence RET3 end of BAC #
DEFINITION      BACR16D16 of RPI1-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION      AL053013
VERSION      AL053013.1 GI:4934461
KEYWORDS      GSS.
SOURCE      fruit fly.
ORGANISM      Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1 (bases 1 to 925)
AUTHORS      Genoscope.
TITLE      Direct Submission
JOURNAL      Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT      Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
```



```
Db      61 GTGGCCCTTTGAGTGTCTATGCCTCCTCATCAACAGCATGGACACTTGTCGTGAGATGNG 120
Oy      1018 CTGCGCTCTTGTACAGCGGTTCTGSAGAAGTACACCAGTAGTGCGGCCAGCCGCTTTTGCCCC 1077
Db      121 GTGGCGCTCTTGTGTGGAGAGCTGTGCTGCACACCATCTGGCTTAAGTTGGCTTGGGCAC 180
Oy      1078 TTGAGACGTCTGCAC 1092
Db      181 CTCGACACTGTTTCCC 195

RESULT   13
CN506XK/c LOCUS DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC # BARLN09 of Rpci-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
AL066051 AL066051.1 GI:4945019 GSS. fruit fly; Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 935) Genoscope. Direct Submission Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr) determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named Rpci-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the Isongenic strain y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm. Location/Qualifiers
1..935 /organism="Drosophila melanogaster" /db_xref="taxon:7227" /clone_lib="Rpci-98" /cloned_in="BARLN09" /note="end : 17"

BASE COUNT    257 a       170 c        162 g         96 t     250 others
ORIGIN
```

OY	980	CGCTACCGGGCAGGAGCGTGTGCGGGAAGCATGTCCTTGCCCTTGCTACGGGGTTCC	1039
Db	755	GSGSCSSSSSBSGCGGCGGCGSSGCSGCCGSSGSGCGSCGCCSCGCCSCGSS	696
OY	1040	TGAGAATGCACCAAGTGGGCGCACCGCGCTTTTGCCCCCTTGAGACTGCTGCACGCGCTAG	1099
Db	695	GCGSGCGCGSGGCGCGCGCCGCCSCCCCSCSSSSSCSSSCSSSSSCSSGCC	636
OY	1100	GAGCGCTCTCCCGCGGGGCGGCTCCAGCCGACTGCGATGTCATTGTACTCTCGGCTCC	1159
Db	635	GCCCCGCGCGCCGCGCCGCCGCCGAGCAGCAGCAGCAGMAGVAGASRSRYVSGSSSS	576
OY	1160	TCTACCCTTAGCGGAGGAGC	1180
Db	575	GSASGCGCGCGMGCRAGSGKM	555
RESULT 14			
CNS017SY/c			
LOCUS			
DEFINITION			
CNS017SY 1101 bp DNA GSS 26-JUL-1999			
Drosophila melanogaster genome survey sequence sp6 end of BAC			
BACN37L08 of DrosBAC library from Drosophila melanogaster (fruit			
fly), genomic survey sequence.			
AL108460			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
fruit fly,			
Drosophila melanogaster			
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;			
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			
1 (bases 1 to 1101)			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Direct Submission			
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :			
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr			
- Web : www.genoscope.cns.fr)			
Determination of this BAC-end sequence was carried out as part of a			
collaboration with The European Drosophila Genome Project (EDGP) -			
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC			
library (Dros BAC) was made by Alain Billaud at CPH (Centre			
d'Etude du Polymorphisme Humain) with funding provided by a MRC			
project grant. The DNA was prepared from embryos by Alain Bucheton			
and Genevieve Payan. It has been constructed in the vector			
pBelobAC11.			
FEATURES			
source			
Location/Qualifiers			
1..1101			
/organism="Drosophila melanogaster"			
/plasmid="pBelobAC11"			
/db_xref="taxon:7227"			
/clone_1lb="DrosBAC"			
/clone="BACN37L08"			
/note="end : SP6"			
BASE COUNT			
254 a 176 c 160 g 152 t 359 others			
ORIGIN			
Query Match 5.3%, Score 63; DB 83; Length 1101;			
Best Local Similarity 15.6%; Pred. No. 0.0045;			
Matches 73; Conservative 206; Mismatches 189; Indels 0; Gaps			
0.0;			
OY	724	TTTCGGAACCGGAGCTTCAGCGCGCGGCTTCATTGTGGCGTGGAGACGAGTGGCCT	783
Db	1073	TTBBSBSTSTSTBTBBSBSSBBBSGSSSSSBSSSTSSSSSTSSSTSSB	1014
OY	784	CCAGCAAACTGTGCTACGCGCTCGGACACTGTGTGTGCGCTCGAGAGGCGGCGCC	843
Db	1013	TBTSTSSSTTTTTTTTTTTTTSTBTBBSBBSBSSSSSTSSBBSKSPBTSSBSSTSS	954
OY	844	GCGCAGGCGACTTGACACCGGTGTTCGCGCGCGCGCTACGCGCTGGGAGCTCGGTCTG	903
Db	953	ASBSSSSSSBBSSTTSTSTBSBBSSTSSSGSSSSSBSTBSBSBHTTTTBTkTSrS	894

Thu Jun 8 15:53:48 2000

us-08-900-220-8.rst

Page 11

THIS PAGE BLANK (u9PT0)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2000, 06:51:49 ; Search time 975.44 Seconds

(without alignments)
6739.859 Million cell updates/sec

Title: US-08-900-220-7

Perfect score: 1622
Sequence: 1 CATCAGCCCAAGAGAGACC.....CATGGAGGCCCATTCCTCC 1622

Scoring table: IDENTITY_NDC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST: *
1: em_est1: *
2: em_est2: *
3: em_est3: *
4: em_est4: *
5: em_est5: *
6: em_est6: *
7: em_est7: *
8: em_est8: *
9: em_est9: *
10: em_est10: *
11: em_est11: *
12: em_est12: *
13: em_est13: *
14: em_est14: *
15: em_est15: *
16: em_est16: *
17: em_est17: *
18: em_est18: *
19: em_est19: *
20: gb_est1: *
21: gb_est2: *
22: gb_est3: *
23: gb_est4: *
24: gb_est5: *
25: gb_est6: *
26: gb_est7: *
27: gb_est8: *
28: gb_est9: *
29: gb_est10: *
30: gb_est11: *
31: gb_est12: *
32: gb_est13: *
33: gb_est14: *
34: gb_est15: *
35: gb_est16: *
36: gb_est17: *
37: gb_est18: *
38: gb_est19: *
39: gb_est20: *
40: gb_est21: *
41: gb_est22: *
42: gb_est23: *
43: gb_est24: *
44: gb_est25: *

45: gb_est26: *
46: gb_est27: *
47: gb_est28: *
48: gb_est29: *
49: gb_est30: *
50: gb_est31: *
51: gb_est32: *
52: em_est20: *
53: em_est21: *
54: em_est22: *
55: em_est23: *
56: em_est24: *
57: em_est25: *
58: em_est26: *
59: gb_est33: *
60: gb_est34: *
61: gb_est35: *
62: gb_est36: *
63: gb_est37: *
64: gb_est38: *
65: em_est27: *
66: em_est28: *
67: em_est29: *
68: em_est30: *
69: gb_est39: *
70: gb_est40: *
71: gb_est41: *
72: gb_est42: *
73: gb_est43: *
74: gb_est44: *
75: em_est31: *
76: em_est32: *
77: em_est33: *
78: em_est34: *
79: gb_est45: *
80: gb_est46: *
81: gb_est47: *
82: gb_gss1: *
83: gb_gss2: *
84: gb_gss3: *
85: gb_gss4: *
86: em_gss1: *
87: em_gss2: *
88: em_gss3: *
89: em_gss4: *
90: gb_gss5: *
91: gb_gss6: *
92: gb_gss7: *
93: gb_gss8: *
94: gb_gss9: *
95: em_gss5: *
96: em_gss6: *
97: em_gss7: *
98: em_gss8: *
99: em_gss9: *
100: em_gss10: *
101: em_gss11: *
102: gb_gss10: *
103: gb_gss11: *
104: em_gss12: *
105: gb_gss13: *
106: gb_gss14: *
107: gb_gss15: *
108: gb_gss16: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result

Query

SUMMARIES

No.	Score	Match	Length	DB	ID	Description
1	352.6	21.9	512	30	AA245525	AA245525 my52c03.r
2	323.2	19.7	452	44	A1325370	A1325370 A134b11.y
3	322.4	19.9	435	51	A1337325	A1337325 z17b12.y
4	252	15.5	324	27	AA032692	AA032692 m13b11.t
5	247.8	15.3	299	69	AA144802	AA144802 EST291875
6	232.8	14.4	464	28	AA099745	AA099745 z17b12.x
7	231.8	14.3	291	81	AA430605	AA430605 70323.MAR
8	199.4	12.3	753	36	AA628967	AA628967 af28h05.s
9	171.2	10.6	456	63	AA1958076	AA1958076 fc9e001.y
10	158	9.7	459	41	AT015155	AT015155 ct7h05.s
11	119.4	7.4	463	49	AT668359	AT668359 mu11g06.x
12	115.6	7.1	537	49	AT645932	AT645932 mu11g06.y
13	81.8	5.0	314	28	AA064660	AA064660 zm13f08.s
14	63.2	3.9	925	82	CNS0091P	AL053013 Drosophill
15	62.2	3.8	288	28	AA076750	AA076750 7B05E11.C
16	58	3.6	925	82	CNS0091P	AL053013 Drosophill
17	56	3.5	932	82	CNS00720	AL066742 Drosophill
18	56	3.5	1101	83	CNS0175Y	AL108460 Drosophill
19	54.4	3.4	510	60	A1810983	A1810983 tu26h12.x
20	54.4	3.4	522	63	AT1963227	AT1963227 wt29b06.x
21	54	3.3	935	82	CNS006XK	AL066051 Drosophill
22	52.8	3.3	1203	83	CNS015Y4	AL106054 Drosophill
23	52.6	3.2	415	23	H57782	H57782 y728e12.s1
24	52.6	3.2	935	82	CNS006XK	AL066051 Drosophill
25	50.8	3.1	1100	83	CNS016KD	AL108655 Drosophill
26	50.8	3.1	1101	83	CNS017SY	AL108460 Drosophill
27	49.6	3.1	932	82	CNS00720	AL066742 Drosophill
28	49.2	3.0	844	82	CNS0052P	AL066552 Drosophill
29	49.2	3.0	910	82	CNS006ON	AL065529 Drosophill
30	47.8	2.9	1201	83	CNS016BR	AL106545 Drosophill
31	47.6	2.9	844	82	CNS0052P	AL056652 Drosophill
32	47.2	2.9	912	82	CNS006N3	AL065775 Drosophill
33	46.8	2.9	1101	82	CNS006G8	AL072149 Drosophill
34	46.8	2.9	1101	83	CNS017V2	AL108535 Drosophill
35	46.2	2.8	839	82	CNS004NB	AL054280 Drosophill
36	46	2.8	782	84	B21685	B21685 F9E2-Sp6.IG
37	45.8	2.8	843	82	CNS00CS1	AL059666 Drosophill
38	45.8	2.8	1101	83	CNS014Y2	AL104756 Drosophill
39	45.4	2.8	744	83	CNS017ZK	AL108659 Drosophill
40	45.4	2.8	744	83	CNS017ZK	AL108659 Drosophill
41	45.2	2.8	964	82	CNS003WG	AL065534 Drosophill
42	44.8	2.8	324	51	AT151714	AT151714 cn1le04.x
43	44.8	2.8	423	72	AW161570	AW161570 au7le05.y
44	44.6	2.7	1009	82	CNS010EW	AL098882 Drosophill
45	44.6	2.7	1203	83	CNS015YR	AL106077 Drosophill

ALIGNMENTS

RESULT	1	AA245525	512 bp	mrna	EST	10-MAR-1997
LOCUS	AA245525	my52c03.r1	Barstead mouse	pooled organs	MPLRB4	Mus musculus CDNA
DEFINITION	AA245525	clone IMAGE:699460 5'	similar to TR:G443944	G443944	INDIAN HBDGHHOG	;; mrna sequence.
ACCESSION	AA245525	AA245525.1	GI:1876329			
VERSION	AA245525	EST				
KEYWORDS	SOURCE	house mouse.				
ORGANISM	house mouse.					
REFERENCE	1	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	1	(bases 1 to 512)				
		Marras, H., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, B., Scheinberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.				
TITLE		The WashU-HMT Mouse EST Project				
JOURNAL		Unpublished (1996)				

COMMENT

On Apr 14, 1993 this sequence version replaced gi:693301.
 Contact: Maria M/Mouse EST Project
 WashU-HMT Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.wustl.edu
 This clone is available royalty-free through LML; contact the
 IMAGE Consortium (Info@image.llnl.gov) for further information.
 MGI:433020
 Seq primer: -28m3 rev2 ET from Amersham
 High quality sequence stop: 511.
 Location/Qualifiers

FEATURES

source
 1..512
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:699460"
 /clone_lib="Barstead mouse pooled organs MPLRB4"
 /sex="mixed"
 /tissue_type="pooled organs"
 /dev_stage="7 day"
 /lab_host="DH10B"
 /note="Organ: pooled. Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site: 1: EcoRI; Site: 2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACGATCGATGAGTGGAGCGCCGCTTTTGTGTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors [GTTGATCGGTAC], digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library constructed by Bob Barstead."
 BASE COUNT
 91 a 161 c 143 g 115 t 2 others
 ORIGIN

Query Match 21.7%; Score 352.6; DB 30; Length 512;
 Best Local Similarity 84.9%; Pred. No. 6.9e-60;
 Matches 417; Conservative 0; Mismatches 71; Indels 3; Gaps 2;

QY	868	CCGCTACCTGCTCTTACGATGACAAATCACAGGAGCGGCGCTCCGGGCA	927
DB	24	CTGCCACCTGCTCTTACGATGAGCAATCATACAGACAGCCATTCGCGCA	83
QY	928	CATTGCGACCGACGTCAGCTGCGCATAGCTGCTGCTGGGTGCGAGCGCTGC	987
DB	84	CATTGCGACCGACGTCAGCAACAGGCAATATGCTGTATCAGGATCAGGCGCTCC	143
QY	988	AGCTTGGCCGCTGTCAGCTGCTCTACACAGTGGCCCTCGGGCTACGCCGCTCA	1047
DB	144	AGCTTGGCCGCTGTCAGCTGCTCTACACAGTGGCCCTCGGGCTACGCCGCTCA	203
QY	1048	CAAGCATGAGACACTGGTGTGAGAGATGTGTGATCTGCTGGCGGCTGGCTG	1107
DB	204	CAAGCATGAGACACTGGTGTGAGAGATGTGTGATCTGCTGGCGGCTGGCTG	263
QY	1108	ACCACCACTGCTAGTGGCTTCTGCGCCCTGAGACTCTTTCACAGCTTGGATGG	1167
DB	264	ACCACCACTGCTAGTGGCTTCTGCGCCCTGAGACTCTTTCACAGCTTGGATGG	321
QY	1168	GCAGCTGAGACCCGGGGGAGGTGTGATGTATCCCACTGCTTACGCCCTGGGGC	1227
DB	322	GCAGCTGAGACCCGGGGGAGGTGTGATGTATCCCACTGCTTACGCCCTGGGGC	381
QY	1228	GCTCTGTTAGAGAGGAGCTTCCACCACTGGGAGTGTCCGGGAGGAGGACTGA	1287
DB	382	NNCTTCTTGAAGAGACACCTTCTCCTGAGCTGTCTGTGGGAGGAGGACTGA	441
QY	1288	AGAGCTC-CACGCTGCTCTCTGAGACTGTACTGTGGTCCAGAGGCTCTCAGCG	1346
DB	442	GGGAGCTTACACAGCTGCTCTGAGAGCTGTCTGTGATCCAAAGGCTCTCAGCG	501
QY	1347	GAGGAGCTGG 1357	

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 291)	Bovine; Bos.			
Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Fairbanking,S.C., Fieking,B.A., Kohler,G.A. and Keeler,J.W.				
Design and use of four pooled tissue normalized cDNA libraries for EST discovery in cattle				
Unpublished (2000)				
On Mar 10, 1998 this sequence version replaced gi:2948930.				
Contact: Smith TPL				
USDA, ARS, US Meat Animal Research Center				
PO Box 166, Clay Center, NE 68933-0166, USA				
Tel: 402 762 4366				
Fax: 402 762 4390				
Email: smith@email.marc.usda.gov				
Single pass sequencing. Bases called and trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 20 and -minmatch 12 options.				
PCR primers				
FORWARD: AGGAACAGCTATGACCAT				
BACKWARD: GTTTCCAGTACAGCAG				
Plate: 37 row: N column: 8				
Seq primer: ATTAGTGACACTATAG.				
Location/Qualifiers				
1..291				
/organism="Bos taurus"				
/db_xref="taxon:9913"				
/clone_11b="MARC 4BOY"				
/issue_type="pooled"				
/lab_host="DH10B"				
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI; Library made from pooled tissue from day 20 and day 40 embryos."				
BASE COUNT	44 a	92 c	96 g	59 t
ORIGIN				
Query Match	14.3%	Score 231.8;	DB 81;	Length 291;
Best Local Similarity	87.3%;	Pred. No. 2.9e-36;		
Matches 254; Conservative	0;	Mismatches 37;	Indels 0;	Gaps 0;
QY 1079	GATGACATCTGCTTGGGGCGGTGGCTGACACACACCTGCTCATGTTGGCTTGCGCC			
Db 1	GGTGGCCCTTGGCTTGGGGCGGTGGCTGACACACACCTGCTCATGTTGGCTTGCGCC			
QY 1139	CCTGAGACTCTTTCACAGCTTGGCATGGGGCAGCTGGACCCCGGGGGAGGCTGTCATG			
Db 61	CCTGAGACTCTTTCACAGCTTGGCATGGGGCAGCTGGACCCCGGGGGAGGCTGTCATG			
QY 1199	GTACCCCAAGCTGCTTACCGCTGGGGCGTCTCCGTAGAGAGAGGCGAGCTTCACACC			
Db 121	GTACCCCAAGCTGCTTACCGCTGGGGCGTCTCTTGGTGGAGAGAGGCGAGCTTCACACC			
QY 1259	ACTGGGCACTGTCGGGGCGAGGAGCTGAAGAAGACTCCACCGCTGCCCTCTGGAAGTCT			
Db 181	ACTGGGCGTGGGGCGGGCGAGGAGCTGAAGAAGATCCTTACTAGCCCTCCAGAACTGCC			
QY 1319	GTACGGGCTCCAGAAAGCTCTGACGACGAGAGGAGGAGTGGCCCTCGAAGGGA			
Db 241	CACAGGTGTCMAAGGCGTCCACACAGAGGAGGACTTGGCTTGGAAAGGA			
RESULT 8				
LOCUS	AA628967	753 bp	mRNA	EST
DEFINITION	af28h06.s1 Soares_total_fetus_NB2HF8_9w Homo sapiens cDNA clone			
ACCESSION	AA628967			
VERSION	AA628967.1	GI:2541354		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 753)	Eutheria; Primates; Catarrhini; Homidae; Homo.			
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Gelsel, G., Jost, S., Kitzman, D., Kucaba, T., Lacey, N., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, R., Steptoe, M., Tan, F., Treisman, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.				
WashU-NCI human EST Project				
Unpublished (1997)				
On Sep 12, 1996 this sequence version replaced g1:107418.				
Contact: Wilson RK				
Washington University School of Medicine				
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108				
Tel: 314 286 1810				
Fax: 314 286 1810				
Email: est@wustl.wustl.edu				
This clone is available royalty-free through LNL; contact the				
IMAGE Consortium (info@image.lnl.gov) for further information.				
Seq primer: -40m13 fwd. ET from Amersham				
High quality sequence stop: 474.				
Location/Qualifiers				
1..753				
/organism="Homo sapiens"				
/db_xref="taxon:9606"				
/clone="IMAGE:1033019"				
/clone_id="Soares.Total-fetus_ND2HF8_9w"				
/dev_stage="8-9 weeks"				
/lab_host="DH10B"				
/note="Vector: pUT73D-Pac (Pharmacia) with a modified				
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA				
was prepared from mRNA obtained from pooled 8-9 week				
(total) fetus material with a Not I - oligo(dT) primer [5'				
TGTTACCAATCTGAAGGGAGGAGGCGCCGCTTAATTTTITTTT 3'].				
Double-stranded cDNA was ligated to Eco RI adaptors				
(Pharmacia), digested with Not I and cloned into the Not I				
and Eco RI sites of the modified pUT73 vector. Library				
went through one round of normalization, and was				
constructed by Bento Soares and M. Fatima Bonaldo. "				
BASE COUNT	190 a	199 c	219 g	145 t
ORIGIN				
Query Match	12.3%	Score 199.4	DB 36	Length 753
Best Local Similarity	96.8%	Pred. No. 8.4e-30		
Matches 214, Conservative 0, Mismatches 6, Indels 1, Gaps 1,				
OY 1402 CCGTCCGATGAAGATACACCATTTAGACTGAGTGGCAACACAGGTCCTCCACCCG 1461				
DB 670 CCGTCCGATGAAGATACACCATTTAGACTGAGTGGCAACACAGGTCCTCCACCCG 611				
OY 1462 GCGTGGTGTAGTCATGAGCTGCAAGCTAGCTGGCGAGGGAGTGGTGGACCCCT 1521				
DB 610 GCGTGGTGTAGTCATGAGCTGCAAGCTAGCTGGCGAGGGAGTGGTGGACCCCT 551				
OY 1522 CTCCTCTAGAGACCTTGAGGTGGCGAGCGCACATCCCACTACAGCTGCTCTACTACGA 1581				
DB 550 CTCCTCTAGAGACCTTGAGGTGGCGAGCGCACATCCCACTACAGCTGCTCTACTACGA 491				
OY 1582 GTTTTCATCTC-TGCCCTCCCATTTGGGAGAGGCCCATTC 1621				
DB 490 GTTTTCATCTC-TGCCCTCCCATTTGGGAGAGGCCCATTC 450				
RESULT 9				
LOCUS A1958076 406 bp mRNA EST 20-AUG-1999				
DEFINITION Icf90e01.y1 zebrafish washu m1mg EST Danio rerio cDNA 5' similar to				
TR:Q92008 Q92008 VHH-1 PRECURSOR ; mRNA sequence.				
ACCESSION A1958076				
VERSION A1958076.1 GI:5750785				
KEYWORDS EST.				
SOURCE zebrafish.				
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;				

OY	195	TTCAGCAGATTACGCCCCCAATGTGCCCGGGAAGACCCTTGCCGGCCAGCGGCATATGA	254
Db	401	TACAAACAATTAGCCCCCAATGTGCCCGGGAAGACCCTTGCCGGCCAGCGGCATATGA	342
OY	235	GCGAAGATCGCTCGCAGCTCCGAGCGCTTCAAGAAGACTCACCCCATAATTAATCCAGAC	314
Db	341	GCGAAGATCGCTCGCAGCTCCGAGCGCTTCAATAGACTCATCCCCCAATTAATCCAGAC	282
RESULT 11			
LOCUS	A1666359	463 bp	mRNA EST 12-MAY-1999
DEFINITION	mul1g06.x1 Soares_thymus_2nbMT Mus musculus cDNA clone IMAGE:639130		
KEYWORDS	3' similar to TR:Q61488 Q61488 DESERT HEDGEHOG HOMOLOG PRECURSOR ;,		
ORGANISM	Mus musculus house mouse.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
AUTHORS	Euthetia; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index unpublished (1997)		
JOURNAL	On May 18, 1998 this sequence version replaced gi:3136856.		
COMMENT	Other ESTs: mul1g06.y1 Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email: Robert.Strausberg@nih.gov This clone is available royalty-free through LUNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information. This clone was previously sequenced on the 5' end only, this new data is from the 3' end Possible reversed clone: similarity on wrong strand High quality sequence stop: 261. Location/Qualifiers		
FEATURES	1..463		
SOURCE	/organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="IMAGE:639130" /clone_idb="Soares_thymus_2nbMT" /sex="male" /tissue-type="Thymus" /dev_stage="4 weeks" /lab_host="DH10B" /note="Vector: pRTT3D-Pac (Pharmacia) with a modified polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGCAAGTGGAGCGCGCGCTTTTCTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTZ19 vector. RNA provided by Dr. Bertrand Jordan. Library went through two rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."		
BASE COUNT	87 a	124 c	138 g 113 t 1 others
ORIGIN			
Query Match	7.4%	Score 119.4:	DB 49; Length 463;
Best Local Similarity	68.0%;	Pred. NO. 3.6e-14;	
Matches 181;	Conservative 0;	Mismatches 82;	Indels 3; Gaps 1;
OY	103	TGCTCTGCTGATGCTGCCCGCGCATCGGGGTGGCGCGGCTGGGTGGTGGCGACCGCC	162
Db	426	TGGCACTCTGTGGCACTATCTGCGCAGAGCTGGCGCGCGCGGACGAGAACCGGTGGCGC	367
OY	163	GCGGACCGGACGCA---AAGTGTGCGCGCTCGCTTACAAAGCATTTACGCCCCCAATGTGC	219

Db 366 GCGCGTAATGTGCGCAGACACTTGTTCCTTCGTCTATCAAGACAGTTTGTGCCAGATATAC 307
 QY 220 CCGAGAAAGACCTTGGGGCCGACGGACCTATGAAAGCAAGATCGCTTCGACGCTCCAGC 279
 Db 306 CCGACCGGACCTTGGGGCCGAGTGGGCGACGGAGGAGGAGGTAACAAGGGGGGTGGAGC 247
 QY 280 GCTTCAAGAGAGCTACCCCAATTCATTCACAGACATCTTCCTCAAGAGCAGAGGAGACA 339
 Db 246 GCTTCCGGAGACTCGTACCCACTACAAACCCGACATATCTTCAAGATGAGGAGACA 187
 QY 340 CAGCGCGCCGACCGCTCATGACCAG 365
 Db 186 GCGCGCAGACCGCCTGATGACAGAG 161

RESULT 12
 AT645932
 LOCUS
 DEFINITION
 5' similar to TR:061488 061488 DESERT HEDGEHOG HOMOLOG PRECURSOR ; ,
 mRNA sequence.
 AT645932
 ACCESSION
 VERSION
 KEYWORDS
 EST.
 AT645932.1 GI:4724407
 SOURCE
 ORGANISM
 house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 537)
 NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 On May 18, 1998 this sequence version replaced gi:3137802.
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert_Strausberg@nih.gov
 This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (infoimage.llnl.gov) for further information.
 This read is a RESEQUENCE of a previously sequenced mouse clone
 This read has been verified (found to hit its original self in the
 correct orientation)
 Possible reversed clone: similarity on wrong strand
 Seq primer: -40RP from GIDCO
 High quality sequence stop: 478.
 Location/Qualifiers
 1..537
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone IMAGE:6391307
 /clone_id="Soares_thymus_2NBMT"
 /sex="male"
 /tissue_type="Thymus"
 /dev_stage="4 weeks"
 /lab_host="DH10B"
 /note="vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was primed with a Not I - 01950(dt) primer [5',
 TGTTCACCAATCTGAAAGTGGGCGCGCGCTTTTATTTTATTTTATTTTATTTT
 3']; double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. RNA
 provided by Dr. Bertrand Jordan. Library went through two
 rounds of normalization, and was constructed by Bento
 Soares and M.Fatima Bonaldo."

BASE COUNT 107 a 159 c 164 g 107 t
 ORIGIN

Query Match 7.1% Score 115.6; DB 49; Length 537;
 Best Local Similarity 67.3%; Pred. No. 21e-13;

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)



Creation date: 02-12-2004
Indexing Officer: ATEKLY - ALEM TEKLAÿ
Team: OIPEScanning
Dossier: 08900220

Legal Date: 04-23-2001

No.	Doccode	Number of pages
1	SRNT	54

Total number of pages: 54

Remarks:

Order of re-scan issued on

